1-1-2000

Authentic assessment of speech-language skills in bilingual students with communicative disorders in an urban school system.

Carey E. McGinn
University of Massachusetts Amherst

AUTHENTIC ASSESSMENT OF SPEECH-LANGUAGE SKILLS IN BILINGUAL STUDENTS WITH COMMUNICATIVE DISORDERS IN AN URBAN SCHOOL SYSTEM

A Dissertation Presented

by

CAREY E. MCGINN

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF EDUCATION

September 2000

School of Education

AUTHENTIC ASSESSMENT OF SPEECH-LANGUAGE SKILLS IN BILINGUAL STUDENTS WITH COMMUNICATIVE DISORDERS IN AN URBAN SCHOOL SYSTEM

A Dissertation Presented

by

CAREY E. MCGINN

Approved as to style and content by:

Kenneth A. Parker, Chair of Committee

Robert Maloy, Member

Maria D. Gonzáles, Member

Sharon Edwards, Member

Bailey W. Jackson, Dean
School of Education

DEDICATION

Especially to my mother, Avis Johnson, whose thirst for learning has always inspired me;

To my family and friends who have supported me and tolerated my absence from so many social occasions;

To my students who show me every day why this was important to me;

To Bobby and Kyle who taught me what this project was really about.

ACKNOWLEDGMENTS

Gratitude is expressed to Dr. Kenneth A. Parker whose guidance, support, and encouragement have sustained me throughout this endeavor.

Appreciation is extended to Dr. Robert W. Maloy who so willingly provided his assistance and expertise in the field of education through all phases of this project.

Appreciation is also gratefully extended to Dr. Maria Diana Gonzales, whose knowledge and expertise in the field of bilingual/bicultural communication disorders has provided such an indelible mark on the development on this dissertation. Thank you on behalf of my students.

Many thanks are accorded to Dr. Sharon A. Edwards whose commitment to her students and joy in their learning has been an inspiration for me (yes, we still have our Writing Boxes!). For your faith in my ability to complete this project, I thank you.

To my students, parents, and colleagues who took time from their busy schedules to openly and candidly answer my interview questions. The sharing of their opinions, experiences, and dreams about this topic has made this a truly collaborative endeavor.

Special appreciation is extended to Linda Ohl whose collaboration and contributions as a colleague have enriched my life and my own learning.

Very special thanks are offered to Mike Westort who has assisted me in innumerable ways from guiding me through the maze of D-forms to editing my work.

To my neighbor, editor, and close friend, Marcia Stratton, whose unfailing patience, sympathetic ear, and wise counsel have guided me throughout this project.

Thanks are also due to Dr. Rebecca Field for her editorial assistance as well as her support.

For AG, thank you for the dream.

ABSTRACT

AUTHENTIC ASSESSMENT OF SPEECH-LANGUAGE SKILLS

IN BILINGUAL STUDENTS WITH COMMUNICATIVE DISORDERS

IN AN URBAN SCHOOL SYSTEM

SEPTEMBER 2000

CAREY E. MCGINN, B.S.,WORCESTER STATE COLLEGE

M.S. WORCESTER STATE COLLEGE

Ed.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Kenneth A. Parker

This study used a case study methodology to describe the implementation of authentic assessment procedures by a bilingual speech-language pathologist in an urban school system over the course of one school year. The specific purposes of the study were to investigate the ways in which authentic assessments could be used to document progress, differentiate a language disorder from a language difference, and determine the language of instruction in bilingual students with communicative disorders.

The initial research strategy involved the collection of authentic assessment data from 65 students in preschool through grade six who received speech-language services or had been referred for an initial evaluation. With the exception of students who had been referred for initial evaluations, authentic assessment procedures were integrated into the delivery of speech and language services. Field notes and reflections on the process of the data collection were completed several times weekly.

Interviews were conducted with 10 parents, 15 teachers, and 13 students about their perspectives on the use of authentic assessment procedures. The three research questions were answered through the inductive analysis of the data drawn from the field notes, reflections, and interviews. The data were analyzed using the constant comparison method.

This study found that a multifaceted approach to communicative assessment that combined standardized measures and authentic assessment techniques facilitated in the differentiation of language disorders from language differences. This comprehensive approach to assessment complied with the mandates of the Individuals with Disabilities Act of 1997 (IDEA-97), decreased the likelihood of assessment bias, and identified specific learning strategies and classroom accommodations that enhanced students' ability to access the regular education curriculum. The implementation of a dual language, multidimensional and curriculum-based assessment model helped to determine the language of instruction in bilingual students with communication disorders. Authentic assessment techniques were found to be an ideal means for documenting students' progress in attaining specific curriculum-based communication goals as required by IDEA-97.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

In this era of high-stakes testing, students enrolled in public schools are participating in an increasing number of standardized assessments that will exert a major influence on the course of their young lives. In many states, tests have been developed that will eventually be used to determine whether or not students receive their high school diploma. The pressure to educate students to the high standards established by these tests has resulted in an increased number of referrals to special education where students are exposed to another regime of assessments (Cummins, 1984). The inappropriate use of biased high-stakes assessments to determine eligibility for services has resulted in the over-representation of culturally and linguistically diverse students in special education programs (Cummins, 1984, 1989; Taylor, 1986, 1993). The improvements in educational outcomes generated by this dependence on high-stakes assessment practices for the next millennium have yet to be determined; however, their impact on current educational practices is unmistakable.

In response to the current over-dependence on standardized tests as the sole means of measuring educational outcomes, a growing number of individuals have begun to reclaim the central role of school-based educators in the assessment process. Authentic assessments, with their long history of use for classroom-based educational decision-making (Wiener & Cohen, 1997), have re-emerged as an alternative assessment option. The evaluation of student achievement or performance through the use of activities that reflect classroom goals, curricula, and or instruction (O'Malley,

1996) has a long tradition in education, however authentic assessment procedures have been less than enthusiastically embraced in the medically based field of communication disorders. Despite the thrust for educationally relevant therapeutic practices, school-based speech-language pathologists (SLPs) continue to rely on standardized tests for decision-making purposes (Secord, Wiig, Damico, & Goodin, 1994). The present study investigates the ways in which authentic assessments can be used to document progress, differentiate language differences from language disorders, and determine the language of instruction in bilingual students with communicative disorders.

Statement of the Problem

With the Massachusetts Educational Reform Act in 1993, statewide broad educational goals, called the Massachusetts Common Core of Learning, were developed with the intent of increasing educational performance for all students. The purpose of these goals was to improve the educational achievement of all pupils including students with disabilities and those in the process of acquiring English as another language. Students previously exempted from participation in state mandated high-stakes assessments because of disabilities or limited English language skills, now take part in this annual testing ritual in third through tenth grades.

Classroom teachers in urban settings are faced with an increasing number of culturally and linguistically diverse students and students in the process of acquiring English as another language. Many feel overwhelmed by the prospect of being held personally accountable for ensuring that all of their students perform well on these

high-stakes assessments. For those students who may be performing below expected levels for a variety of reasons, a referral to special education may be initiated in order to alleviate the teacher's responsibility for the anticipated poor performance on the high-stakes assessment.

SLPs frequently serve as "gatekeepers" for special education services and are often the first ones who are approached by teachers when there is concern about the academic skills of culturally and linguistically diverse students. The high-stakes assessments administered by SLPs are often the key link to the chain of referrals that lead to categorical diagnosis and special education placements that can result in lowered expectations, and risk for academic failure for culturally and linguistically diverse students (Quinn, Goldstein, & Pena, 1996). In many cases, SLPs do not possess the requisite training for the conduction of non-biased assessments of students from culturally and linguistically diverse backgrounds. In 1994, a national survey of public school clinicians was conducted. Approximately 76% of the respondents had no previous coursework or classes that addressed services with multicultural populations. Ninety percent of the respondents could not communicate in a second language with sufficient fluency for the conduction of an assessment (Roseberry-McKibbin & Eicholtz, 1994).

SLPs, untrained to distinguish language differences from language disorders, administer standardized tests that have not been normed on culturally and linguistically diverse populations. Due to the inherent biases of these instruments, culturally and linguistically diverse students often score poorly on these tests and are presumed to be learning and/or language disabled. The placement of these students

into special education programs at rates higher than their representation in local school populations is overwhelmingly due to the inappropriate use of standardized tests that are biased against individuals from culturally and linguistically diverse populations (Cummins, 1984).

The over-representation of culturally and linguistically diverse students in specific categories of special needs illustrates one of the consequences of these high-stakes assessment practices (Cummins, 1984, 1989; Hamayan & Damico, 1991; Kayser, 1995, 1998; Taylor, 1993). In many urban areas in the United States, culturally and linguistically diverse students are over-represented in disability categories relative to their proportion in the total population. Nationwide, Latinos represent 8.2% of the population, yet 56.8% are classified as learning disabled, 23.7% are classified as speech-language impaired, and 23.7% are classified as mentally retarded (Office for Civil Rights, 1993).

The consequences of high-stakes assessments are especially grim for bilingual students with language disorders. A large body of research validates the critical connection between language competence, literacy, and academic success (Apel, 1999; Catts, 1997, 1999; Catts & Kamhi, 1986, 1999; Kamhi & Catts, 1986; Synder & Downey, 1991: Wallach & Butler, 1994; Wiig & Semel, 1984). Given the linguistic basis for many reading difficulties, a large proportion of these students may be slow to develop literacy skills. Educators, cognizant of the need for the student to do well on state mandated assessments, may seek assistance for the anticipated poor performance on these high-stakes tests. Placement in a more restrictive special education program

in which the instruction is delivered in English is the prescribed solution to the problem.

It is a common misconception that bilingual language/learning disabled students will be confused by dual language instruction. It is assumed that this will result in difficulty in the acquisition of literacy skills and the inevitably poor performance on high-stakes tests. The research literature supports the hypothesis that first language instruction facilitates skill development in both languages for bilingual students with special needs (Bruck, 1978, 1982, 1985a, 1985b; Gutierrez-Clellen, 1999; Kiernan & Swisher, 1990; Maldonado, 1994; Ortiz, 1984; Perozzi, 1985; Perozzi & Chavez-Sanchez, 1992; Thordardottir, Weismer, & Smith, 1997). Other studies demonstrate that the level of second language proficiency that a bilingual child attains is partially a function of the level of competence that he/she has developed in the first language at the time when extensive exposure to the second language begins (Anderson, 1989; Cummins 1979, 1981, 1986; Shutnabb-Kangas, 1984). Despite the evidence that supports first language instruction for bilingual students with communicative disorders, many educators continue to believe that it is in the best interests of the child to choose one language of instruction. The language of choice is usually English – the language of the majority culture.

This monolingual instructional model has the potential to create an educational environment in which the bilingual child receives instruction in a foreign-language without the continued support of the first language, leaving the child without a foundation for acquiring the second language (Shutnabb-Kangas & Toukomaa, 1976). Given the high stakes consequences of state mandated tests, the determination of the

most appropriate language of instruction is a critical decision for bilingual students with communicative disorders.

It is ironic that this proliferation of high-stakes testing practices is occurring at a time in which there is an increasing level of dissatisfaction with the use of norm-referenced tests in the field of special education (Hamayan & Damico, 1991; Kayser, 1998; Secord, et al., 1994). A growing body of research supports the use of authentic assessment procedures including portfolios, curriculum-based, performance-based, and dynamic assessment procedures as alternatives to norm-referenced tests, particularly in urban areas with culturally and linguistically diverse populations (Hamayan & Damico, 1991; Kayser, 1998; Larson & McKinley, 1995; Nelson, 1993, 1994, 1998; Pena, 1993, 1996; Schraeder, Quinn, Stockman, & Miller, 1999). Not only do authentic assessments provide for a less biased assessment of a student's performance level on specific tasks or components of the curriculum, information gained from the assessment can be used to guide instructional practices within the educational setting.

The specific purposes of this study are to:

(1) Identify how authentic assessment practices can be used to determine the language of instruction for bilingual students with communicative disorders.

(2) Identify how authentic assessments can be utilized to document progress in the attainment of speech-language therapy goals as outlined in a student's Individual Educational Program (IEP).

(3) Identify how authentic assessments can be used to supplement standardized assessment procedures for the differentiation of language differences from language disorders.

Significance

It is projected that by the year 2000, one out of every three people in the United States will be a person of color (American Speech-Language-Hearing Association, 1991). Speech-language pathologists who work in the school setting can expect that one third of their caseloads will be comprised of culturally and linguistically diverse students (Cole, 1989). Norm-referenced tests can no longer be used as the sole means for identifying students with speech and language disorders. SLPs need to develop appropriate means of accurately assessing these students in a non-biased manner. Authentic assessments provide a viable alternative to norm-referenced tests for SLPs who work with culturally and linguistically diverse populations in educational settings.

Two major problems result from the inappropriate use of norm-referenced tests with students from culturally and linguistically diverse populations. The first problem is that the cultural and linguistic experiences of these students may be very different from those for which the test was standardized. Test norms are most reliable for individuals who are most like the subjects used in the test standardization population, and least reliable for individuals who are least like the standardization sample population (Huang, Hopkins, & Nippold, 1997). The utilization of norm-referenced tests with culturally and linguistically diverse students is likely to produce unreliable

results when used for the differentiation of language disorders from language differences (Damico, 1991; Huang, Hopkins, & Nippold, 1997; Kayser, 1989; Taylor & Payne, 1983). The indiscriminate use of standardized instruments has the potential to produce biased test results that do not represent the linguistic capabilities of the students and often inaccurately label them as language and/or learning disabled. Secondly, students may not have had sufficient exposure to the types of tests used in this country and may lack the test taking skills necessary to perform well on these measures. If test taking is viewed as a complex social activity (Mehan, Hertwreck, & Meihls, 1986) that is learned within a cultural context, many students from culturally and linguistically diverse backgrounds have not been socialized to the activity of test taking. The rigid standardized testing procedures that do not allow deviations from a prescribed script and/or offer a limited options for "correct" responses, may seriously underestimate students' capabilities (McLaughlin, Blanchard, & Osanai, 1995).

A number of researchers have advocated the use of authentic classroom-based communicative assessment procedures (Fradd & McGee, 1994; Gillam & McFadden, 1994; Launer, 1998; Larson & McKinley, 1995; Nelson, 1989, 1994; Prelock, 1997; Secord, et al., 1994). Authentic assessment refers to those activities that "reflect the actual learning and instructional activities of the classroom and out-of-school worlds" (Hiebert, Valencia, & Afflerbach, 1994, p. 11) of the students. Thus, assessors take into consideration the social and cultural milieu of the student both within and outside of the classroom setting. The linkage of standardized and authentic assessments offers the potential for decreasing the biases generated by the exclusive use of norm-referenced tests.

The Individuals with Disabilities Acts of 1997 (IDEA-97; Pub. L. 105-17, U.S.C. Sess. H. R. 5, 1997) mandates that speech-language assessments be educationally relevant and reflect a student's current performance level within the general education curriculum. Parents must be informed of their child's progress toward the attainment of individualized goals on a quarterly basis. In addition, the student's likelihood of meeting the annual goals identified in the IEP must be determined each quarter based upon the student's current rate of progress. If it is unlikely that a student's goals will be achieved, a new IEP must be generated. School-based SLPs question their ability to realistically comply with this mandate given their large caseloads (J. O'Toole, personal communication, September 2, 1998).

Typically, SLPs determine each student's progress in therapy by completing formalized testing on a yearly basis. The testing completed for the purpose of generating new objectives for the IEP is a time consuming endeavor, requiring the cancellation of numerous therapy sessions. Because norm-referenced tests can be quickly and easily administered, SLPs frequently rely on these instruments to measure progress in therapy. However, norm-referenced tests are generally not sensitive enough to document changes resulting from speech-language intervention (McCauley & Swisher, 1984). In addition, norm-referenced tests are completed individually under artificially constructed communicative situations. The results obtained from these assessments do not reflect a student's communicative performance within the general education curriculum.

Authentic assessment procedures focus on a student's progress toward the mastery of target goals in an ongoing teaching/learning cycle (Wiener & Cohen,

1997). These procedures use multiple means to measure a student's day-to-day performance in the classroom or therapy settings. Authentic assessments can serve as structured and systematic measures of language behaviors within meaningful, context-bound classroom activities (Notari-Syverson & Losardo, 1996). Because assessment is frequently integrated into instructional or therapeutic activities, the dilemma of canceling therapy in order to complete testing is minimized. Authentic assessments offer an ideal means for documenting progress toward the mastery of specific educationally-based communication goals as required by IDEA-97.

The model of bilingual education in operation in Massachusetts mandates that students achieve sufficient language skills for transitioning into the monolingual program within three years time. Language dominance tests have been developed for the purpose of evaluating a student's proficiency in each language (Mattes & Omark, 1984) and for determining whether a student possesses the requisite English skills for transitioning into the monolingual program. Critics argue that language dominance tests do not measure a bilingual student's communicative abilities in each language in meaningful speaking contexts (Mattes & Omark, 1984). Since the majority of these tests focus on surface structures such as grammatical usage, researchers have questioned the ability of these tests to measure the higher-level language skills necessary for success in the monolingual classroom (Cummins, 1981, 1984). Because authentic assessments focus on a student's performance using classroom goals or curricula, they may offer a viable option to language dominance tests for the accurate evaluation of the cognitive/academic language proficiency skills required for success in the monolingual program.

A large body of research is highly critical of the use of high-stakes assessments for use with students from culturally and linguistically diverse backgrounds, however a limited number of options have been offered for school-based SLPs with few resources and very limited time constraints. The present study seeks to fulfill this void by identifying how authentic assessment procedures can be used for three purposes: (1) to select the language of instruction for bilingual students with communicative disorders, (2) to document progress in speech-language therapy, and (3) to supplement standardized assessments for the differentiation of language differences from language disorders.

Limitations of the Study

The findings derived from this study were limited to 65 monolingual Spanish-speaking or bilingual Spanish/English-speaking students referred for speech-language assessment or diagnosed with a communicative disorder. The majority of these students were enrolled in bilingual classrooms in an urban setting in Massachusetts. The findings from this study may not be generalized to students enrolled in bilingual programs in other communities in Massachusetts, to speakers of other languages, or to students without communicative disorders.

This study was also limited by the researcher's role, as evaluator/participant in the process of organizing and managing the implementation of authentic assessments with bilingual students, the majority of whom were enrolled in speech-language services. Although students, parents, and educational personnel were interviewed about their viewpoints on authentic assessment procedures, the investigator's

perspective was central to this study and cannot be considered as representative of school-based SLPs. Because the purpose of this study was to initiate a transformation in assessment practices by a bilingual SLP within a specific educational setting, the results are limited to the context of the study and may not be generalized to other educational settings or systems (Cochran-Smith & Lytle, 1993).

The findings from this study depended upon data collected in interviews with parents and educational personnel. The reliability and validity of the collected data were influenced by the ability and/or willingness of the interview subjects to take time away from their classrooms and other responsibilities to participate in these interviews. The internal validity of this portion of the study was further influenced by the interviewees' willingness to explore the topic of authentic assessments as well as their willingness to express their opinions openly and candidly (Bogdan & Biklen, 1992; Spradley, 1979).

The reliability and validity of the data collected from the parent and student interviews were influenced by the dynamics of the student, parent, and SLP relationship. Students may have been reluctant to freely express their opinions with an individual who has power over them (Seidman, 1998). In turn, parents may have been hesitant to express their true opinions with someone who was invested in the research project and had power over their children. Finally, colleagues may have been reluctant to express their beliefs candidly due to concern that this may have placed them in a vulnerable or awkward position and/or may have altered their working relationship with the investigator, especially if their opinions were not supportive of the research topic (Seidman, 1998). On the contrary, the close relationship between

the researcher, the students, and their parents may have resulted in a fuller and more open expression of interviewees' opinions and beliefs.

Definition of Terms

Authentic Assessments	Procedures used to evaluate a student's achievement or performance through the use of activities that reflect classroom goals, curricula, and instruction or real-life performance (O'Malley, 1996).
Bilingual	The ability to use two languages with varying degrees of proficiency and in a variety of contexts, e.g., understanding, speaking, reading, and writing (Grosjean, 1982).
Communicative Disorder	An impairment in the ability to receive, process, understand and send concepts of verbal, nonverbal and written symbol systems that significantly affect academic, social-behavioral, or vocational areas (Nicolosi, Harryman, & Kresheck, 1978).
Cultural/Linguistic Diversity	Individuals whose language and cultural learning origins differ from those of the mainstream (Weismantel & Fradd, 1989).
High-Stakes Assessment	The testing of students for purposes such as grade level retention or advancement, high school graduation, and/or eligibility for special programs

	or services (National Clearinghouse for Bilingual Education, 1997).
Individual Educational Program (IEP)	A written service plan that indicates a student's current level of performance, the educational services the student is to receive, the long and short term instructional goals for the provided services, and a specified time period in which the services will be provided (Flower, 1984). An IEP was referred to as an Individual Educational Plan prior to IDEA-97, but IEP refers to Individual Education Program throughout this paper.
Language Difference	A variation of a symbol system used by a group of individuals that reflects regional, social, or cultural/ethnic factors (Taylor, 1986). A language difference is not synonymous with a speech or language disorder.
Language Disorder/Deficit	Difficulty in the ability to understand or use linguistic symbols for communicative purposes as dictated by the norms of the community (Taylor, 1986).
Metacognition	The conscious awareness of the thinking process, including perceiving, remembering, imagining,

	conceiving, judging, and reasoning (Nelson, 1998; Nicolosi et al., 1978).
Metalinguistic abilities	The process of consciously reflecting on language, including the simultaneous processing of language at more than one level, and the knowledge of the specific labels used to talk about language (Nelson, 1998).

CHAPTER 2

REVIEW OF THE LITERATURE

Introduction

The purpose of assessment in the field of communication disorders will be summarized before reviewing the literature on the current assessment options used with students from culturally and linguistically diverse backgrounds. The primary goals are to examine the applicability of these assessment options for the following purposes: (1) to determine the language of instruction in bilingual students with communicative disorders, (2) to document progress in speech-language therapy, and (3) to distinguish language differences from language disorders. The components addressed in the literature review reflecting the above goals are: (1) the definition of the role of assessment in communication disorders, (2) the identification of the role of norm-referenced assessments with students from culturally and linguistically diverse populations, (3) the identification of the role of criterion-referenced assessments with students from culturally and linguistically diverse populations, (4) the identification of dynamic assessment procedures, (5) the identification of authentic assessment procedures, and (6) the identification of additional sources of information.

Assessments in the Field of Communication Disorders

A major responsibility for school-based SLPs is to select and complete appropriate assessments followed by the thoughtful evaluation of the collected data

(American Speech-Language-Hearing Association, 1998). Although the terms "testing", "assessment", and "evaluation" are often used interchangeably, these words actually represent distinct activities. Testing refers to the use of specific tests or defined test procedures for the generation of a score or rating of an individual (Taylor & Payne, 1983). Assessment refers to "data collection and the gathering of evidence", while evaluation "implies bringing meaning to the data through interpretation" (Routman, 1994, p 302).

The American Speech-Language-Hearing Association (ASHA, 1998) defines the characteristics of an appropriate speech-language assessment. First, it provides answers for the same developmental questions that are asked by language research (e.g., What are the important hallmarks of the developing system?). Second, an appropriate speech-language assessment accommodates and describes the wide array of individual differences between students. Third, rather than focusing exclusively on an individual's deficits, it also elicits evidence of a student's abilities and strengths. Fourth, an appropriate speech-language assessment provides information that can be used for educational planning and goal setting. Fifth, it describes with-in child and across-child changes over time (Fischgrund, 1996, as cited in ASHA, 1998). Finally, an appropriate speech-language assessment yields baseline documentation for progress monitoring.

Ideally, the diagnosis of a speech-language disability in a monolingual English-speaking child consists of the gathering of information from various sources (e.g., teachers and parents), the completion of a number of norm-referenced, standardized tests, and the observation of the child in one or more settings. When the

child in question is from a culturally and linguistically different background or does not speak English, the assessment process becomes more complex (Kayser, 1989, 1995, 1998; Langdon & Saenz, 1996; Taylor, 1986, 1993).

According to the Education of All Handicapped Children Act (PL 94-142, 1975), assessment and evaluation procedures must be conducted in the language or mode of communication in which the child is most proficient and must be administered in a nondiscriminatory manner. Taylor and Payne (1983) define nonbiased assessment as a data collection process in which the assessment and evaluation are conducted using instruments and procedures that discriminate only in those areas for which they were designed (normal versus disordered communication) and do not discriminate against an individual because of cultural or social variations or due to other factors such as gender, socioeconomic status, and dialect use.

The following section will explore the use of norm-referenced testing procedures with culturally and linguistically diverse students. This will be followed by a discussion of the use of criterion-referenced, dynamic, and authentic assessment procedures for reducing bias in communicative assessments, documenting progress in speech-language therapy, and for determining the language of instruction for bilingual students with communicative disorders. This section will end with an examination of the pivotal role of social/educational history and level of acculturation for the accurate evaluation of the collected data.

Norm-Referenced Assessments

Norm-referenced tests sample the end products of behavior during a single point in time and report the results in terms of a standardization group (Haywood, Brown, & Wingenfeld, 1990). Tasks that are used to assess current levels of functioning in norm-referenced tests are predetermined, decontextualized, and product oriented (Pena, 1993). These tests are administered using standardized procedures. Students are ranked and compared to other children of a similar age. Variations in performance are considered to be attributable to variations in ability (McCauley & Swisher, 1984). Norm-referenced language tests are assumed to provide an adequate and valid estimate of an individual's language skills and abilities (Damico, 1991; McCauley & Swisher, 1984). In the field of communication disorders, norm-referenced tests are often positively viewed by clinicians. They are considered to provide a means of assessing the efficacy of clinical services in a quantifiable way and of selecting and sorting children for educational placement (Gutierrez-Clellen, 1996).

The use of norm-referenced tests in the assessment of culturally and linguistically diverse students is a politically sensitive issue (Mattes & Omark, 1984). The placement of children from cultural and linguistic minority groups into speech-language therapy and special education classes at rates higher than their representation in the local population is considered to be the result of inappropriate assessment procedures (Taylor, 1986). Placement decisions are often based upon the results of norm-referenced tests that are presumed to accurately identify the presence of a language disability and to predict future performance (Plante & Vance, 1994). "That standardized tests are objective, neutral, and somewhat independent of contextual

influences that may affect the measurement of language behavior is incompatible with a perspective that recognizes language variations and sociocultural diversity within and across speech communities" (Gutierrez-Clellen, 1996, p. 31).

Norm-referenced tests assume homogeneous exposure to the content of test items (Figueroa & Garcia, 1994) as well as linguistic, cultural, and background experiences in the student being tested. For example, Gutierrez-Clellen and Iglesias (1987, as cited in Lidz & Pena, 1996) found that the Puerto Rican mothers in their study tended to use commands, deixis (e.g., esto/this), and functional descriptions of objects (e.g., you sit on it, go like this + hammering motion) during interactions with their pre-school children. Upon entering a school setting, many of these children may be assumed to have poor vocabulary skills when they evidence these behaviors during naming tasks. The reliability of norm-referenced tests used in isolation for the assessment of language abilities in culturally and linguistically diverse students is highly questionable.

Attempts have been made to reduce the inherent biases of norm-referenced tests, however the outcomes of these alternations and adaptations are doubtful (Vaughn-Cooke, 1983). For example, it has been suggested that existing tests should be standardized on non-mainstream populations. This suggestion assumes that non-mainstream populations are homogeneous and does not take into consideration the variation of cultural and linguistic background experiences of different populations. Vaughn-Cooke argues that this practice would result in lower norms for non-mainstream English-speakers and eventual comparisons between racial groups that would contribute to lowered expectations for minority children.

A second suggestion for reducing bias in norm-referenced tests has been to include a small percentage of minorities in the standardization sample when developing the test. This alternation violates the assumption of homogeneity of the normative sample that is a critical assumption of test standardization (Pena, 1993). This practice does not increase the assessment validity of the instrument (Weiner & Hoock, 1973) and could lead to ranking groups in terms of presumed ability which would be detrimental to culturally and linguistically diverse students (Cole & Means, 1981; Wolf, Bixby, Glenn, & Gardner, 1991).

A third suggestion has been to develop local norms for the targeted norm-referenced test (Langdon, 1989). In addition to being very time-consuming and possibly expensive, this practice does not take into consideration the following: (1) variations within populations, (2) mismatch of test construct, and (3) low expectations of certain groups (Pena, 1993). Intragroup heterogeneity diminishes the feasibility of this option (Duran, 1989; Gutierrez-Clellen, 1996). A variety of factors including first and second language proficiency, cultural background, socioeconomic level, background experiences, and level of acculturation must be taken into consideration when developing local norms. In addition, if the test construct does not match the student's background experiences, the instrument becomes a test of acculturation (Pena, 1993). The development of local norms does not improve the construction of a test with questionable validity for a target population.

A number of tests have been translated into other languages in order to assess the knowledge of specific content. This creates a number of problems when these instruments are used for diagnostic and placement decisions. Direct translations do

not yield equivalent forms, as some words do not have exact counterparts in other languages (Mattes & Omark, 1984; Pena, 1993, 1996). For example, arroz/rice may refer to a tomato-based and spicy food for Spanish connotations versus a white or brown substance for Anglos (Kayser, 1998). The level of difficulty of a particular item, word, or task may be affected (Figueroa, 1989). Direct translations do not alter the culturally bound content of the test (Olmedo, 1981) nor increase a student's familiarity with the culture of test taking. Duran (1989) argues that " ... every test requires that examinees understand specialized uses of language for the sake of test taking itself" (p. 54). Educational decisions based upon information derived from translations of norm-referenced tests are highly questionable as the test and testing situation continue to contain both technical and situational biases (Pena, 1993).

Language dominance tests are norm-referenced tests whose purpose is to determine a student's linguistic proficiency in each language. Language proficiency refers to the level of skill or degree of control that an individual exhibits over a language (Payan, 1984).

Cummins (1984) distinguished two types of language proficiency that he labeled as basic interpersonal communicative skills (BICS) and cognitive/academic language proficiency skills (CALPS). Basic interpersonal communicative skills refer to the language proficiency required for communication within daily-living types of contexts, while cognitive/academic language skills require the manipulation of language within decontextualized academic situations. Basic interpersonal language skills are more typical of the everyday language that is required for social purposes outside of the classroom (e.g., communication with peers at recess time).

Communication is embedded within the context of the situation and meaning can be actively negotiated by the participants. Because cognitive/academic language skills are not supported by the context of the situation, correct interpretation of the message relies primarily on linguistic cues and knowledge of the language itself. Cognitive/academic language skills are typically associated with the linguistic demands of the classroom setting (e.g., understanding the social studies curriculum) as well as the content of high-stakes assessments. Students who are learning English as another language typically acquire basic interpersonal communication skills in approximately two or three years time, while at least five to seven years are required for the acquisition of English cognitive/academic language proficiency skills (Collier, 1987; Cummins, 1981, 1984).

A number of researchers have questioned the adequacy of language dominance tests to evaluate a student's ability to function successfully in the monolingual learning environment as they do not measure cognitive/academic language skills (Cummins, 1981, 1984; Lopez, Lamar, & Scully-Demartini, 1997; Ochoa, Galarza, & Gonzalez, 1996; Ortiz & Polyzoi, 1988). Other authors have argued that educational decisions based upon information derived from language dominance tests should be made with caution and have recommended the development of authentic language tasks (Zehler, Hopstock, Fleischman, & Greniuk, 1994).

Criterion-Referenced Assessments

Criterion-referenced tests are a frequently cited alternative to norm-referenced tests. In contrast to norm-referenced assessments that compare a student's

performance to a norm or standardized sample, criterion-referenced tests measure a student's mastery of specific objectives that have been defined by predetermined standards of judgement or criterion (Notari-Syverson & Losardo, 1996; Secord, et al., 1994). Criterion-referenced assessments can be used to: (1) validate norm-referenced test results, (2) document progress or assess educational outcomes, and (3) select targets for intervention and educational objectives (Secord et al., 1994).

When applied to the field of communication disorders, criterion-referenced assessments use one or more focused probes that target a particular communication skill, rule, or strategy. The items within the probe may examine the same skill at different levels of difficulty or within different contexts (e.g., Is the student able to predict outcomes in four different hypothetical problem-solving situations?). Criterion-referenced probes can be utilized to assess specific language skills that are reflected within the educational, vocational, or social-communicative settings (Secord et al., 1994; Nelson, 1993, 1994).

An advantage of criterion-referenced assessments is that they can be directly linked with classroom curriculum and can provide information that is applicable to goal setting and intervention. Because criterion-referenced assessments are not conducted using standardized testing procedures, adaptations and accommodations are permissible as the goal of the assessment is to elicit a student's responses under optimal conditions (Notari-Syverson & Losardo, 1996). As a result of the flexibility of these assessment procedures, criterion-referenced tests can be used with students of varying age levels as well as in a variety of target languages.

A disadvantage is that a number of criterion-referenced assessments have included isolated items from norm-referenced tests into their testing protocol. This practice limits the educational value of the test as it invalidates the age-equivalency scores for the selected items (Johnson, 1982). Pena (1993) criticizes the passive role of the student in the assessment process as well as the assumption that performance reflects ability. Criterion-referenced assessments attempt to fragment language and learning into isolated components that are easily measured. The attempt to document progress in the attainment of specific goals of intervention (e.g., goals on an IEP) can result in part-to-whole instructional techniques.

Dynamic Assessment Procedures

One viable alternative to norm-referenced tests is the implementation of dynamic assessment procedures (Feuerstein, 1979; Feuerstein, Rand, Haywood, Kyram, & Hoffman, 1995; Lidz & Pena, 1996; Pena, Quinn, & Iglesias, 1992). Several researchers have advocated the use of dynamic assessment procedures as a way of reducing bias in psychological testing (Figueroa, 1989; Feuerstein et al., 1995; Lidz, 1987; Sewell, 1987). Dynamic assessment uses a test-teach-retest model. Baseline data are collected during the initial testing phase. This information guides the procedures attempted during the teaching phase (Pena, 1996). Some practitioners utilize intervention that consists of graduated prompts (Campione & Brown, 1985), while others use a mediated learning experience (Feuerstein, 1975, Feuerstein et al., 1995; Lidz, 1987, 1991).

Mediation is defined as the purposeful teaching of skills and strategies for problem solving by the adult mediator (Lidz & Thomas, 1987). Incoming information is mediated or modified by the adult so that the child can understand it in a meaningful way (Tzuriel & Klein, 1987). The emphasis is on an individual's learning process and potential, rather than on the products of past learning (Haywood, Brown, & Wingenfeld, 1990). During the mediation phase of dynamic assessment, the examiner analyzes the amount and type of input needed to produce a change in the individual being tested (Pena, 1996). During the re-test phase, the examiner is able to observe how the student applies strategies learned during the mediation phase for the solution of novel problems. Dynamic assessment connects evaluation procedures with interventions by providing information about how the learner responds to the interventions that were utilized during the mediation phase.

One of the unique features of dynamic assessment is the interactive nature of the assessment process (Haywood, et al., 1990). The examiner responds to the child in specific ways during the assessment process in order to modify his/her responses to different tasks. Dynamic assessment assumes that all children are capable of learning and it is the responsibility of the examiner to discover ways to modify the child's learning. If a change is not observed, it is because the examiner has not been sufficiently diligent in the investigation of those modifications that would produce a change (M. Samuels, personal communication, May 15, 1996). The examiner's role shifts from that of a passive observer to that of an interactive teacher/learner (Haywood et al., 1990; Lidz, 1987, 1991).

The application of dynamic assessment procedures to the field of communication disorders is emerging (Acevedo, Hardee, & Fernandez, 1997; Butler, 1997; Gillam & McFadden, 1994; Launer, 1998; Lidz & Pena, 1996; Pena, 1993, 1996; Pena, Quinn, & Iglesias, 1992). Because dynamic assessment measures an individual's learning potential, based upon previous experiences and the provision of new learning opportunities, it is considered to be an inherently nonbiased assessment procedure. It also holds great potential as an assessment procedure that can distinguish language differences from language disorders. It can also assist in the selection of the language of instruction for bilingual students with communicative disorders, e.g., which language required a lower level of mediation and less effort to learn a new skill, strategy, or accomplish a task (Acevedo, Hardee, & Fernandez, 1997).

Authentic Assessment Procedures

In response to the plethora of concerns with standardized testing, an array of alternative assessment procedures has been developed. These assessment procedures have been referred to by a variety of terms. The most commonly cited terminology in the field of alternative measurement includes informal, performance-based, and authentic assessments. Although all of these alternative procedures evaluate the quality of a student's work over an extended time period (Mitchell, 1992), these terms are not necessarily synonymous. Alternative assessments share two core features: (1) they are regarded as alternatives to standardized assessments, and (2) they directly

examine an individual's performance on tasks that are significant and important for daily living purposes (Worthen, 1993).

Informal assessment techniques refer to "the structured and systematic observation of behaviors within meaningful, context-bound situations in multiple settings" (Notari-Syverson & Losardo, 1996, p. 259). These techniques can be incorporated into classroom routines and learning activities and do not interfere with instructional time (Navarrete, Wilde, Nelson, Martinez, & Hargett, 1990). These authors describe informal assessment techniques as being timely as opposed to time consuming, representative of curricular goals, and meaningful for both the teacher and the student. In contrast to standardized assessments, informal assessments provide continuous, ongoing documentation of a student's progress. The information gained from informal assessments contributes to the effective planning of instruction and goal setting. Navarrete and his colleagues recommend that informal assessments be used to supplement results derived from standardized tests, especially when working with students from culturally and linguistically diverse backgrounds.

Performance-based assessments typically require students to "construct a response, create a product, or demonstrate applications of knowledge" (O'Malley, 1996, p. 239). Students demonstrate their mastery of the curriculum by providing solutions to presented problems through the manipulation of materials or the completion of hands-on activities. Oral reports, writing samples, group and/or individual projects, exhibitions, and demonstrations are examples of activities utilized during performance-based assessments.

Performance-based assessments provide evidence of a student's communicative and academic skills in relation to process-oriented educational standards (Navarrete & Gustke, 1996), rather than the exclusively product-oriented standards offered by norm-referenced tests. They link instruction with learning and ensure meaningful and multiple assessment opportunities. These assessments provide educators with the opportunity for the measurement of student progress, as well as for the documentation of the educational process. Time management for scoring and reporting data is one of the major challenges associated with performance-based assessments (Navarrete & Gustke, 1996).

Authentic assessment refers to "procedures for evaluating student achievement or performance using activities that represent classroom goals, curricula, and instruction or real-life performance (O'Malley, 1996, p. 237). Authentic assessment uses techniques that more realistically and fully describe how a student completes a complex task that represents daily activities. It translates educational objectives into activities that require the application of a variety of language and learning skills that cross subject areas and disciplines, thus preparing students with the cognitive skills necessary to function as literate adults (Gillam & McFadden, 1994; Leshe & Jett-Simpson, 1997; Wiener & Cohen, 1997). Authentic assessments provide students with an array of tasks that reflect the priorities and challenges of exemplary educational activities: conducting research, collaborating with others on a project, and/or writing, revising, and discussing a written piece of work (Wiggins, 1990).

Ecological authentic assessments are ongoing and integrated with instruction and learning (Leshe & Jett-Simpson, 1997). They comprise an integral component of

the ecosystem of the learning environment (Bartoli & Botel, 1988) and involve the collection of representative samples of authentic learning as it occurs as students engage in authentic activities. In contrast to norm-referenced assessments that measure speech and language skills at one point in time, ecological procedures require multiple assessments in multiple contexts over time. This ongoing approach to assessment is an ideal means for documenting progress in the acquisition of literacy, language, and learning skills (Leshe & Jett-Simpson, 1997).

An essential component of authentic assessment is the active participation of students in the evaluation process. Self-assessment contributes to students' direct involvement in the learning process and the integration of cognitive abilities with motivation and a positive attitude toward learning (O'Malley, 1996). It assists students in understanding the connection between personal effort and successful learning thus encouraging the assumption of responsibility for directing their own learning. The active engagement of students in the analysis and evaluation of their own learning is more likely to provide students with a self-empowered educational experience, an essential ingredient for combating institutional racism and the reversal of the high failure rate of culturally and linguistically diverse students (Cummins, 1986).

Curriculum-Based Language Assessments

Curriculum-based language assessments use components of the classroom curriculum to evaluate a student's linguistic knowledge, skills, and strategies relative to what the student must do to be successful within the classroom or school setting

(Nelson, 1989, 1992, 1993, 1994, 1998). The purpose and function of curriculum-based language assessment is distinct from more traditional assessments in several ways.

First, it is a criterion-referenced procedure that provides the opportunity to determine if a student can meet the criterion for success within a given classroom setting. This approach is distinct from norm-referenced tests that measure a student's ability to perform within normal limits on a given test (Nelson, 1992). A second difference is that curriculum-based language assessments examine the processes of learning, not the products of learning. A third distinction between traditional testing and curriculum-based language assessment is that it evaluates the dynamics of the child within the learning context. The student's language needs are examined in relation to the communicative demands within the school setting as opposed to traditional assessments that look at language outside of the actual learning environment. Finally, curriculum-based language assessment uses interview and observational data as a basis for the selection of assessment contexts and content (Nelson, 1992).

Specific contexts within the curriculum that are described as being problematic for the student are identified as "zones of significance" (Nelson, 1993, 1994). These are highly personalized contexts identified through ethnographic interviewing procedures with the student, teacher, and parent as being important for the student. The information from these interviews is supplemented with classroom observations of the curricular demands and of the student interacting with the curriculum (Nelson, 1993).

Curriculum-based language assessments examines the interaction between the language learner and the curricular contexts from both an *outside-in* and *inside-out* perspective (Nelson, 1994). The *outside-in* look identifies the communicative and linguistic demands of a given curricular activity and considers ways to modify them so that the curriculum is more accessible for the student. The *inside-out* look identifies a student's current level of communicative knowledge, skills, and abilities and examines ways to increase the effectiveness of their implementation through a variety of mediational processes and other intervention strategies.

A major advantage of evaluations that incorporate curriculum-based language approaches is that assessments are ongoing and progress can be measured through functional outcomes (Nelson, 1994). Both quantitative and qualitative documentation can be used with curriculum-based language assessment.

A second advantage of curriculum-based language assessment is the reliance on materials from a student's actual classroom curriculum for data collection and the documentation of progress (Leshe & Jett-Simpson, 1997; Nelson, 1998). A student's gradually increased ability to access the regular educational curriculum can be measured over time through curriculum-based language assessment, a process that cannot be documented through the use of traditional standardized assessments.

The Role of Observation in Authentic Assessments

Observation plays a central role in the conduction of any assessment activity. Norm-referenced, curriculum-based, dynamic, and authentic assessments all require the utilization of observational and coding abilities. Although norm-referenced tests

do require observation skills, the highly standardized and rigid structure of this type of testing limits the role of observation (Secord, et al., 1994).

"Observation is the foundation of the cycle of assessment-diagnosis-intervention in communication disorders" (Secord et al., 1994, p. 49). Accurate and systematic observations provide the SLP with information about the student's abilities and areas of need within the context of the classroom environment.

Effective observations allow the observer to collect ongoing contextually embedded data during actual communicative events within a variety of settings, thus increasing the authenticity of this approach to assessment (Secord et al., 1994). Systematic and purposeful observations within the classroom setting also provide information about the underlying patterns and organization of the classroom context as well as insight into teacher expectations and the student's ability to meet these expectations (Nelson, 1994). Observations facilitate in the effective planning of classroom or curricular accommodations that will help students to access the curriculum. The systematic conduction of observations also assists in the accurate documentation of a student's growth and progress in both the classroom and therapy settings.

A variety of structured observational tools are available for use by SLPs. Silliman and Wilkinson (1991) classify these tools into three categories: categorical, narrative, and descriptive.

Categorical

Categorical tools are closed systems that have predetermined categories for the coding of events and behaviors noted during the observation (Silliman & Wilkinson, 1991). These structured observational tools quantify and/or summarize the collected data in a numerical fashion, such as an overall rating, a summary score, a ranking, a frequency count, or a percentage. Checklists, protocols, and rating scales are examples of categorical tools.

Running records and miscue analyses are categorical tools that examine a student's reading skills while reading aloud. A running record codes, scores, and analyzes precise reading behaviors for the purposes of detecting reading difficulties early and monitoring a student's reading progress (Clay, 1993; Allington & Cunningham, 1996). Educators can use a running record to observe the reader's implementation of specific reading behaviors such as word analysis strategies, self-corrections, fluency, word accuracy, and word meaning strategies during any oral reading event (Clay, 1993). Reading comprehension is often determined based upon the student's answers to specific questions, story retelling, and/or observation of the reader's responses during story discussions (Jett-Simpson & Leslie, 1997). A running record provides quantitative as well as qualitative information about a student's interaction with a text, can be used with beginning readers as well as more skilled readers, and can be adapted to a variety of languages (Escamilla, Andrade, Barsurto, & Ruiz, 1996). Because the student's actual reading of a text is used, a running record mirrors the learning activities of the classroom. In contrast to standardized assessments that are an indirect means of measuring a student's learning, a running

record is a resource that provides direct and specific information about each student's learning (Clay, 1993).

A miscue analysis is another type of reading assessment that examines a student's use of decoding strategies, reading strategies, and comprehension while reading aloud (O'Malley, 1996). It is based upon the belief that reading is an active language process that is influenced by the knowledge that readers have about language (Goodman, Watson, & Burke, 1987). A miscue analysis demonstrates how readers use language-based cueing systems to actively construct meaning from a written text. These systems include graphophonemic cues (the relationship between sounds and the written form of language), syntactic cues (the interrelationship of words, sentences, and paragraphs within a written text), semantic cues (the meaning of language), and pragmatic cues (the social-cultural context). Readers' deviations from the written text or miscues are not regarded as errors, but as a valuable tool for providing insight into the active process of meaning construction from the text (Gillam & Carlile, 1997; Gillam & McFadden, 1994; Goodman, Watson, & Burke, 1987). Following the oral reading of the text, comprehension is assessed through an unaided story retelling procedure. Follow-up questions may be used to expand upon the retelling, to elicit specific information about crucial aspects of the story, and/or to stimulate the reader to reflect upon the passage (Weaver, 1994).

Weaver (1994) provides an alternative to the analysis of miscues developed by Goodman, Watson, and Burke (1987). She suggests that miscues be analyzed via the following five questions: (1) Did the miscue go with the preceding text? (2) Did the miscue go with the following text? (3) Did the miscue preserve essential meaning? (4)

Was the miscue self-corrected? (5) Was the miscue either meaning preserving or self-corrected (referring to questions 3 and 4).

Y. Goodman and K. Goodman characterize miscues as the "windows on language processes at work" (1998, p.122). The analysis of miscues provides information about language and learning disabled students' ability to integrate information from several cognitive-linguistic systems (graphophonemic, syntactic, semantic, and pragmatic) while constructing meaning during a reading aloud activity. For example, miscues evidenced by the reader during an oral reading task may be consistent with those demonstrated during oral communication, suggesting a language system that is not fully developed in all contexts of the curriculum (Nelson, 1994).

Gillam and Carlile (1997) examined the differences in oral reading and story retelling between students with specific language impairments (SLI) and typically achieving students using a miscue analysis approach. The students with SLI evidenced a significantly higher percentage of miscues than their reading-matched peers and a greater number of their narratives were rated as confusing and incomplete. The authors speculated that the SLI students' lack of prior knowledge about the topics of the stories combined with their slow language processing skills and/or memory deficits may have accounted for the differences found between the two groups.

A number of studies have also investigated the correlation between reading and listening comprehension at different grade levels (Sticht, Beck, Hauke, Kleiman, & James, 1974; Sticht & James, 1984). These researchers found that the correlation between reading and listening increased from first through sixth grade. According to Snow, Burns, and Griffin (1998), a high correlation between reading and listening

comprehension occurs after a child has learned to decode; however, the gap between a student's reading and listening comprehension skills may be quite large, even when the correlation between the two skills is strong. Other investigators have concluded that as skilled readers recognize words from text, they simultaneously employ their language skills (semantic, syntactic, morphologic, and pragmatic) and background knowledge to derive meaning from the text. Therefore, once they can automatically recognize words in print, the processes of reading and listening comprehension become nearly identical (Snow, Scarborough, & Burns, 1999).

A narrative sample is a third type of categorical tool. This procedure examines a student's comprehension skills and knowledge of story structure by analyzing the ability to retell a story that has been viewed (e.g., a movie), listened to, or read by the student (Gutierrez-Clellen, 1998; Morrow, 1988; O'Malley, 1996). Narratives are closely related to classroom activities, particularly reading comprehension and the composition of written stories (Gillam & Carlile, 1997; Silliman & Wilkinson, 1991; Westby, 1994; Westby, Von Dongen, & Maggart, 1989). One author referred to narratives as the bridge between the interactive oral style of interpersonal conversational skills and the cognitive/academic language skills required for literary forms (Westby, 1992). As well as serving as an ecologically valid means for determining a student's understanding of relationships in the world, Westby characterized narratives as a means of assessing the ability to use language to organize this knowledge in a cohesive and coherent manner.

Because the production of narratives reflect the influence of both contextual and cultural factors, extreme caution must be exercised when attempting to distinguish

language differences from language disorders (Gutierrez-Clellen, 1998; Gutierrez-Clellen & Heinrichs-Ramos, 1993; Gutierrez-Clellen, Pena, & Quinn, 1993; Gutierrez-Clellen & Quinn, 1993). In contrast to topic associated narratives exhibited by students from culturally and linguistically diverse backgrounds, the narratives of students with language disorders demonstrate their difficulty with perspective taking, organization, and planning (Silliman & Wilkinson, 1991). A dynamic approach to narrative assessment that acknowledges differences in cultural experiences, exposure to narrative tasks, and assumptions about audience involvement is suggested by some researchers (Gutierrez-Clellen, Pena, & Quinn, 1995; Gutierrez-Clellen & Quinn 1993)

An advantage of categorical tools is that they are easy to administer and low in cost. A second advantage is the flexibility of use for the observation of a wide range of behaviors or individuals. A disadvantage is that they typically do not provide sufficient detail for the qualification of target behaviors. If adequate training is not provided for the observer, the reliability of categorical tools may be compromised. Categorical tools are not recommended for the measurement of the outcomes of specific interventions or when the precise description of a communicative skill is required (Silliman & Wilkinson, 1991).

Narrative

Narrative tools are systematic and detailed written descriptions of observed behaviors that are typically recorded in the same order in which they occurred (Silliman & Wilkinson, 1991). The observation period is variable, ranging from the

recording of a single incident to observations that last a day or longer. Specific examples of narrative tools include running records, anecdotal notes about critical incidents or target behaviors, journal entries, or ethnographic notes of observations compiled during participant observations.

Because narrative tools are easily recorded and written in everyday language, the collected data are easily accessible to a variable audience. The ongoing nature of the recorded observations allows the observer frequent opportunities to review the information in order to discern patterns of communicative behaviors. A major disadvantage of narrative descriptions is that the collection and analysis of data may be influenced by subjective interpretations and biases (Silliman & Wilkinson, 1991). Narrative tools can be used in isolation or in combination with categorical and/or descriptive tools.

Descriptive

Descriptive tools are the verbatim transcriptions of the actual language used during the observation. They typically include detailed descriptions of the observed behaviors as well as the context in which the behaviors were observed. Descriptive tools are frequently used for the identification of developing processes and/or patterns of behaviors observed in different settings (Silliman & Wilkinson, 1991). Because descriptive tools provide a detailed recording of language use and accompanying behaviors, they provide critical information that may not be noted during more casual observations. Technological tools such as video and/or tape recordings are generally required during the data collection phase.

A language sample is a type of descriptive tool. The systematic collection and analysis of a student's verbal output is considered to be a cornerstone of any communicative assessment protocol (Evans, 1996; Nicolosi, Harryman, & Kresheck, 1978). It is usually recommended that the language sample consist of 100 or more utterances collected within three different contexts (Kayser, 1998). Restrepo (1998) suggested the inclusion of a story retelling activity. In her investigation of the identifiers of language impairment in Spanish-speaking children, she found that the number of errors per utterance and an explicit parent interview yielded a discriminant accuracy greater than 90% between Spanish-speaking children with language differences and language disorders.

A number of additional formats have been suggested by researchers for the collection of a representative language sample. Anderson (1996) developed a series of structured language elicitation activities that evaluated preschool children's productive knowledge of grammatical forms through the manipulation of toys and objects. Stockman (1996) introduced an alternative language sampling procedure that incorporated the concept of a minimal competency core while maintaining cultural sensitivity; however, she cautioned against the exclusive use of a language sample for the identification of language disorders in linguistic minority children. While play-oriented activities are often used with preschool children (Linder, 1993), Evans and Craig found an interview format to be a reliable and valid means of collecting a language sample from school-aged children (1992). Nelson (1993) suggested that the interviewing process be supplemented with questions designed to elicit an emotional response from the student.

Language samples for younger children are typically analyzed using a mean length of utterance format; however, several researchers have suggested the use of a mean length of response (MLR) analysis system for Spanish-speaking children (Anderson, 1995; Kayser & Restrepo, 1995; Restrepo, 1998; Restrepo & Gutierrez-Clellen, 1997). This entails dividing the language sample into utterances, counting the number of words contained in each utterance, and then averaging the number of words by the total number of utterances. The results of the MLR are then compared to developmental data reported for Spanish-speaking students from monolingual and bilingual environments (e.g., Anderson, 1995; Linares & Sanders, 1977; Merino, 1992).

For older students, the analysis of terminal units (TU) is advocated as a means for determining the complexity of verbal output (Hunt, 1965), especially when working with Spanish speaking populations (Gillam, Pena, & Miller, 1999; Gutierrez-Clellen & Hofstetter, 1994; Gutierrez-Clellen, Restrepo, Bedore, Pena, & Anderson, 2000; Kayser, 1998; Kayser & Restrepo, 1995; Restrepo, 1998). According to Hunt (1965), a TU contains a main clause and all of its subordinate clauses and modifiers. Gutierrez-Clellen and Hofstetter (1994) adapted Hunt's analysis of TU for Spanish, recommending that cojoined subjectless sentences be counted as a separate TU, rather than the one unit as counted by Hunt. The researchers noted that this adaptation was necessary because in Spanish the verb carries the information about the subject, therefore verbs without subjects are permissible (e.g., corrieron - they ran).

As with the MLR, the average number of TU per utterance is counted and averaged. The level of subordination is calculated by adding the number of clauses

per TU and then averaged. Restrepo (1998) reported that the analysis of TU may be useful for distinguishing language differences from language disorders in bilingual children by determining if both languages evidence a significant number of errors per TU (more than 0.18 errors per TU in each language).

Descriptive tools provide comprehensive information about a student's current level of communication as well as the contexts that support or prohibit its continued development. These tools provide valuable information that can guide instruction and document progress toward the mastery of educational goals (Silliman & Wilkinson, 1991). Training in the accurate collection and analysis of data is typically required. The process of data collection, transcription, and analysis can be a very time consuming endeavor.

Portfolio Assessments

Portfolios have been recently adopted from the field of language arts by SLPs (Gillam & McFadden, 1994; Kayser, 1998; Kratcoski, 1998; McLaughlin, Blanchard, & Osanai, 1995; Roseberry-McKibbin, 1995; Secord, et al., 1994; Wiig & Secord, 1991). Portfolios contain an array of materials that have been systematically collected over time and that reflect the work of an individual student (Secord, et al., 1994). They provide a number of benefits when used in the assessment of a student's communicative skills. Because a portfolio consists of a collection of classroom artifacts, the SLP is able to obtain samples of a student's language abilities within authentic contexts and for authentic purposes. Portfolios also contribute to the

collaborative process by involving the student, teacher, and parents as contributors to assessment (Kratcoski, 1998).

The composition of communication portfolios is variable. They may contain formal and informal work samples (e.g., compositions), audio and/or videotaped recordings of student projects, work, or class projects as well as homework assignments. Observational reports by teachers and parents, anecdotal records, lists of books read, charts or graphs of work completed over a given time period, and lists of activities in which the student has participated may be filed in portfolios. In addition, SLPs may want to collect language samples, narrative samples, journal or learning log entries, peer evaluations, and teacher, parent, and/or student interviews. Testing data, conference notes, and classroom checklists, rating scales, and/or protocols may also be included in portfolios (Kratcoski, 1998; Secord et al., 1994).

The contents of portfolios need to be appropriately and systematically analyzed. This ongoing analysis promotes the generation of accurate conclusions about a student's language abilities or progress in therapy. The contents of the portfolio can be analyzed in a number of ways. For example, work samples can be analyzed descriptively by recording a student's use of specific skills or strategies that demonstrate changes in skill level across samples and over time. Checklists, rating scales, or rubrics (Pierce & O'Malley, 1992) facilitate in the evaluation of written work, oral presentations, or a student's implementation of learning strategies.

Holistic scoring procedures, in which the quality of a specific item is given a quantitative rating (Gillam & McFadden, 1994; Kratcoski, 1998; Pierce & O'Malley, 1992), are frequently used in portfolio assessment. As students with language

disabilities may exhibit inconsistencies in their developmental profiles (Westman, 1990), the holistic scoring of portfolio samples may not provide an accurate assessment of their performance level. Consequently, focused holistic scoring procedures are recommended for use with students with communicative disorders (Semel, Wiig, & Secord, 1989). Specific areas of focus for SLPs may include: (1) the level of abstraction and detail in spoken language, (2) the ability to recall facts or details from a story, (3) the use of linguistic conventions (e.g., word use, word ordering, or social language use), or (4) the mechanics of spoken language (e.g., articulation, rate, fluency, or intonation) (Secord, et al., 1994).

In contrast to norm-referenced tests that provide a "snapshot" of a student's abilities at one point in time, portfolios demonstrate a student's growth, achievement of curriculum requirements, and acquisition of learning strategies through the display of curriculum-based artifacts. Portfolios allow SLPs and other educators to present evidence of a student's strengths and areas of need for functional, authentic, communicative purposes. Portfolios reflect a philosophy that validates the process of active learning and promotes collaboration among educators, students, and families (Sumner, 1991). Through the use of portfolios, a story can be told about the student's progressive development throughout the school year that is accessible to parents, educators, and most importantly the child (A. McPartland, personal communication, November 19, 1998).

In this era of high-stakes testing, portfolios offer an authentic means of evaluating student progress and contribute to thoughtful, educational decision-making and program planning. Portfolios can also be used to compare and contrast a student's

communicative and literacy skills across languages for the accurate and authentic assessment of the most appropriate language of instruction.

Additional Sources of Information

One of the best tools for obtaining a nonbiased assessment is a SLP who is trained in the differentiation of language disorders and language differences, understands the inherent difficulties in the assessment process, and is capable of analyzing data from a variety of sources, including standardized tests, alternative assessments, and student histories (ASHA, 1998). The goal is to conduct a *balanced* assessment so that the school-based SLP can "evaluate the information gained from all assessment data and make an informed decision about eligibility or placement and subsequent intervention strategies" (ASHA, 1998, p. 29). Two factors that facilitate the completion of a balanced assessment are a student's level of acculturation and social/educational history.

The phenomenon of acculturation has a long history of research in the field of anthropology and more recently in the fields of sociology, psychology, and psychiatry (Cuellar, Arnold, & Maldonado, 1995; Cuellar, Harris, & Jasso, 1980; Montgomery, 1992; Padilla, 1980). In the educational setting, a student's level of acculturation can have a significant impact on academic, social, and communicative behaviors (Damico & Hamayan, 1992). Test taking behaviors are also influenced by an individual's level of acculturation (Westby, 1996). SLPs must take a student's level of acculturation into consideration when differentiating language differences from language disorders as well as when determining the most appropriate language of instruction for a

bilingual child with a communicative disorder. The process of acculturation and its influence on educational outcomes will be examined in the next section.

Acculturation and Acculturation Scales

Acculturation refers to a dynamic phenomena that occurs when two autonomous cultural groups come into continuous contact with one another that leads to fundamental changes in one or both cultures (Berry, 1980). The definition implies that acculturation involves fundamental changes at both the macro (social/group) and micro (individual) levels (Cuellar, Arnold, & Maldonado, 1995). When acculturation occurs at the micro level, it is often referred to as psychological acculturation and involves changes in attitude, behaviors, beliefs, and values in the individual (Graves, 1967).

A by-product of the process of acculturation is an increased level of stress and pressure that arises out of the conflict between the individual's first culture and the second culture (Born, 1970). This conflict results in acculturation stress and may be evidenced by behaviors or experiences that may be mildly pathological or disruptive to the individual (Berry, 1980). These behaviors or experiences may include atypical social interactions, psychosomatic illnesses, or feelings of marginality. Williams and Berry (1994) found that acculturation stress could lead to academic difficulties in culturally and linguistically diverse children.

A number of studies have investigated the effects of acculturation on a student's classroom performance (Cloud, 1991; Collier, 1988; Hoover & Collier, 1986; Olmedo, 1980). Research has suggested that acculturation stress can predispose

culturally and linguistic diverse students to negative outcomes such as academic failure (Chavez, Moran, Reid, & Lopez, 1997; Williams & Berry, 1994). In a number of cases, psychological responses to the acculturation process may be misinterpreted as learning problems, language disorders, or other special needs (Adler, 1981; Collier, 1988; Gavillan-Torres, 1984). Exhibited behaviors may include atypical social interactions, confusion regarding locus of control, decreased expressive language skills, and behaviors associated with an increased level of anxiety (Anderson, 1989).

Collier (1988) studied the relationship between the level of acculturation and referral to special education services for culturally and linguistically diverse students. She found no statistically significant difference in achievement test scores between students referred and not referred for special education. The referred and non-referred groups did not differ significantly on their educational profiles, but did differ on their acculturation profiles.

Anderson (1989) examined the relationship between acculturation and the adaptive behavior characteristics of culturally and linguistically diverse students as a means of predicting special education placement. She found differences between regular and special education students for several variables of acculturation. Students in her study who were placed in special education were less proficient in Spanish, had language skills that were less well developed, and had lived in the United States for a longer time period than those who were not enrolled in a special education program.

A wide variety of acculturation scales are presently available (Cuellar, Arnold, & Maldonado, 1995; Cuellar, Harris, & Jasso, 1980; Marin, Otero-Sabogal, & Perez-Stable, 1987; Martinez, Norman, & Delaney, 1984; Olmedo, Martinez, & Martinez,

1978). Two researchers argue that acculturation scales serve as a means of determining the influence of acculturation on the results of psychoeducational assessments (Barona & Miller, 1994). These scales have recently been adapted to the field of communication disorders for determining the language of therapy with bilingual children (Acevedo, Hardee, & Fernandez, 1997) and for selecting the first language to test with bilingual students (Gonzales & Kayser, 1997).

Barona and Miller (1994) developed the Short Acculturation Scale for Hispanic Youth (SASH-Y) for use with students in grades three through eight. An exploratory factor analysis was conducted and found that 80.5% of the explained variance consisted of questions related to language preference when alone or outside of the family context. A second factor accounted for 11.5% of explained variance and consisted of questions related to language usage patterns with family members. Questions related to social relation preferences accounted for 7.9% of the explained variance. Examination of these dimensions led the authors to suggest that "language use in general relates to both extrafamilial and intrafamilial contexts" (p. 160). They believe that the finding that cultural behaviors are contextually related has important implications for psychoeducational assessments and the interpretation of test data.

Barona and Miller (1994) urge evaluators to analyze psychoeducational test results within the context of a student's level of acculturation. This recommendation has applications to the field of communication disorders where the judicious use of acculturation scales can facilitate in the differentiation of language differences from language disorders for culturally and linguistically diverse students. When used in conjunction with the authentic assessment of students' language and literacy skills,

acculturation scales have the potential to improve the rigor of the decision making process for the selection of the most appropriate language of instruction in bilingual students with communicative disorders.

Social and Educational History

This section will explore the importance of analyzing all test results within the context of a student's social and educational history. The importance of this sociocultural perspective and its role in the selection of the most appropriate language of instruction and in the differentiation of language differences from language disorders will be examined.

The compilation of case history information is a component of all speech-language assessments. However, specialized information must be elicited and a cross cultural perspective must be maintained when the student is from a culturally and linguistically diverse background. It is imperative that the data collected during the social/educational history interview be comprehensive in its scope. This information will assist in the differentiation of a language difference from a language disorder, will ultimately determine the type of supported education recommended for the student, and will aid in the selection of the most appropriate language of instruction.

The student's social and educational history should be elicited through a personal interview that seeks to understand the events and information from the informant's perspective (Spradley, 1979). Westby (1990) recommended the adoption of ethnographic interviewing techniques in order "to understand the social situation in

which the families exist and how the families perceive, feel about, and understand these situations" (p. 105).

Questions regarding a student's social and family background supplement the information derived from the completion of the acculturation scale. Information should be obtained about the birthplace of the student, parents, and ideally the grandparents as well as their educational, occupational, and socioeconomic levels (Gonzales & Kayser, 1997; Langdon, 1992). Other topics to be explored during the interview include the family's ethnic identity, the level of contact with the country of origin and plans to return to their homeland (Gonzales & Kayser, 1997; Langdon, 1992).

In addition to the standard questions about the student's communicative development (e.g., at what age did the child say the first word?), information regarding the language usage patterns of the home, school, and neighborhood should be explored (Acevedo, Hardee, & Fernandez, 1997; Gonzales & Kayser, 1997). Inquiries should be made about the family's perceptions of the child's speech and language development and views on disabilities. Information should be solicited about cultural expectations and perceptions regarding communication (e.g., the conversational status of children) as well as the student's exposure to literacy within the home setting (Acevedo, Hardee, & Fernandez, 1997; Gonzales & Kayser, 1997; Langdon, 1992; Westby, 1994).

A comprehensive exploration of the student's educational history should also be completed. In addition to questions about the type of programs in which the student was previously enrolled (bilingual versus monolingual program), specific

information about the language used for instruction should be obtained. Langdon (1992) suggests that queries be made about the student's attendance and any disruptions in education as well as the provision of classroom accommodations and/or enrollment in any specialized program (e.g., Title I, special education). For students who may have attended school in another country, information should be gained about the typical number of students enrolled in a classroom, the length of the school day/year, the level of teacher training, the instructional methods typically used by the teachers, and the age at which students are expected to enter school.

The information obtained from the student's social/educational history provides a backdrop for a sociocultural interpretation of the data obtained during the assessment. It also facilitates in the identification of a language disorder versus a language difference and assists in the determination of the need for supported education services as well as in the selection of the most appropriate language of instruction.

Summary

Given the pervasive influence of high-stakes tests on current educational practices, the assessment approaches used by SLPs must reflect a multi-faceted curriculum-based perspective that is appropriate for utilization with students from culturally and linguistically diverse populations. Authentic assessments focus on a student's ability to construct meaning within the context of classroom-based curricular activities (an *inside-out* approach) as well as examine the sociocultural climate of the classroom and the communicative demands of the curriculum (an *outside-in*

perspective). When used in conjunction with more traditional assessment protocols, authentic assessments hold great promise for limiting bias when attempting to distinguish language differences from language disorders as well as for determining the language of instruction for bilingual students with communicative disorders. The collaborative involvement of families, educators, and students combined with the sociocultural viewpoint provided by the social/educational history and level of acculturation, serve to increase the accuracy and reliability of the evaluation outcomes as well as to decrease assessment biases. Authentic assessment approaches measure a student's ability to construct meaning as a listener, speaker, reader, writer, and thinker (Wells, 1986) on an ongoing basis using curriculum-based language activities. Thus the systematic recording of observations facilitate the documentation of progress toward mastery of target behaviors as required by IDEA-97.

Gutierrez-Clellen (1996) urged SLPs to assess language from an interactionist and constructionist perspective that views language ability and disability within a sociocultural perspective. SLPs can no longer engage in biased testing practices that involve the use of norm-referenced tools in isolation, but must consider all assessments as cultural events that require the assumption of an ethnographic perspective (Taylor, 1993). This approach requires a radical move away from reductionist testing practices conducted in an insular manner, toward the incorporation of authentic assessment techniques that examine a student's ability to use language to successfully access the school curriculum. Given the eventual consequences of high-stakes assessments such as the Massachusetts Comprehensive System (MCAS), culturally and linguistically diverse students need to assume roles as test-takers that

empower, rather than disable them. School-based SLPs must assume roles that empower their students by advocating for the incorporation of socioculturally appropriate authentic instruments into their assessment protocols.

CHAPTER 3

DESIGN AND PROCEDURES

Introduction

These are challenging times when educators are striving to implement new learning standards and prepare students for success on high-stakes assessments (e.g., MCAS). A balanced approach to assessment is crucial (ASHA, 1998). Norm-referenced tests, used in isolation with culturally and linguistically diverse populations, do not accurately differentiate language differences from language disorders (Cummins, 1984, 1986; Gutierrez-Clellen, 1996; Hamayan & Damico, 1991; Taylor, 1993). The majority of norm-referenced tests cannot be used to document progress in speech-language therapy (McCauley & Swisher, 1984), as required by IDEA-97. Commercially available language dominance tests typically assess discrete language skills such as grammar use, but do not accurately measure the cognitive/academic language skills (Cummins, 1984; Lopez, Lamar, & Scully-Demartini, 1997; Ochoa, Galarza, & Gonzalez, 1996) required for success in a monolingual classroom as well as on high-stakes assessments. Given these drawbacks associated with traditional standardized tests, assessment practices that circumvent these areas of need must be implemented. Authentic assessment practices fulfill this need.

This qualitative study examined the use of authentic assessment procedures with bilingual students with communicative disorders. The specific purposes of the study were to:

(1) Identify how authentic assessments could be used to determine the recommended language of instruction for bilingual students with communicative disorders.

(2) Identify how authentic assessment could be utilized to document progress in the attainment of speech-language therapy goals as outlined in a student's IEP.

(3) Identify how authentic assessments could be used to supplement standardized assessment procedures for the differentiation of language differences from language disorders.

Design of the Study

This study used a case study methodology to describe the implementation of authentic assessment procedures by a bilingual SLP in an urban school system over the course of one school year. The goal was to examine the evolution in assessment practices as a consequence of a SLP's involvement with the process of educational change. The case history approach is regarded as an especially appropriate methodology for portraying what occurs when educators transform their educational practice, because it allows other educators a first hand account of the factors most salient to this transformation (Bissex & Bullock, 1987; Edwards, 1997).

This study was heavily influenced by the teacher-as-researcher approach that employs methodical intentional inquiry to produce social change from "the inside out, from the bottom up, through changes in teachers themselves" (Bullock, 1987, p. 27). Teacher research seeks to improve the quality of teaching and learning in schools

through systematic investigation of school-based educational problems. A goal of teacher research is to reform, not replicate, educational practices by "teaching against the grain" (Cochran-Smith, 1991, p. 279). Since the purpose of this present study was to transform the assessment practices used by a school-based SLP with culturally and linguistically diverse students, the teacher-as-researcher model was considered to be most appropriate.

Components of ethnographic research were also incorporated into this research study. Knapp (1979) argues that ethnographic research initially assumes an open-ended exploratory approach to the research problem. This is followed by the researcher's intensive involvement in the social setting being studied, as both an observer and participant. Although multiple research approaches are utilized by ethnographic researchers, key informant interviewing and participant observation techniques are stressed. These techniques provide a means for understanding events from the perspective held by those within the social setting and emphasize the key role that context plays in interpreting events and behaviors (Bogdan & Biklen, 1992, 1998; Knapp, 1979; Maxwell, 1990; Miriam, 1998; Spradley, 1979).

Setting of the Study

This study took place in an urban school system in Massachusetts with a student population of 25,000 enrolled in pre-school through post-graduate programs (technical-training programs). Based upon federal surveys, 51% of the school system's students were considered to be low income, 25% were described as limited English proficient, and 0.06% of the students were enrolled in bilingual programs.

According to the October 1998 census, 45% of the total student population were designated as racial/ethnic minorities with 5% of Native American origins, 10% of African-American origins, 7.1% of Asian origins, and 27.6% of Latino/Hispanic origins. The total special education enrollment was 17.7% with 45.3% of the students designated as racial/ethnic minorities (Office of the Deputy Superintendent, personal communication, November 30, 1998).

The process of authentic assessment implementation was explored in the course of conducting speech-language evaluations with 65 students in pre-school through grade six (see pages 60-63 for specific examples of authentic assessment techniques). Sixty-two percent of these students were enrolled in bilingual classes and spoke Spanish as a first language. The bulk of these students were of Puerto Rican descent; however, a number of these children had origins in the Dominican Republic and El Salvador. Although approximately one half of these students evidenced emerging English cognitive/academic language skills (Cummins, 1984), the majority were in the process of learning English as a second language. Of these students, 40 were enrolled in speech-language services and 16 students also received resource room or inclusionary education services (i.e., students participate in the regular education program for the majority of the school day and receive additional special education supportive services).

Data were also collected from school-aged students referred for an initial speech-language assessment. Of these initial referrals, data were collected from 12 students who were enrolled in the bilingual program. The bilingual SLP was responsible for the determination of the student's eligibility for speech-language

services as well as the language or languages in which the student was to be tested. The completion of record reviews, student/teacher interviews, and an acculturation screening revealed that all of these students were in the early stages of English language learning. Assessments were therefore conducted in Spanish.

Data were also gathered from 12 students who were enrolled in the monolingual program, were suspected of having a communication disorder, and for whom a formal request for a bilingual speech-language evaluation had been initiated. As the school-based SLP is usually responsible for completing the English portion of the testing; these assessments were conducted primarily in Spanish with English used as a facilitating device (Gonzales & Kayser, 1997). Record reviews, interviews with the student, parents, and/or teachers, and the determination of students' level of acculturation supplemented the results obtained from the more formalized assessment procedures. The information derived from these sources was used to differentiate those students who exhibited difficulty because of language loss (the weakening or loss of an individual's first language secondary to intensive exposure to a second language, Schiff-Myers, 1992) and/or lack of experience, from those with a language disorder.

Finally, data were gathered from 13 students referred for an initial assessment through the pre-school arenas that was completed two times monthly with a bilingual psychologist. These referrals were initiated by early intervention programs, pediatricians, social workers, day care centers, Head Start Programs, or the child's family. Although these assessments were frequently play based (Linder, 1993; Westby, 1980), early literacy skills were a component of the evaluation (van Kleek,

1990, 1998). Students were assessed via play-based language probes, observation, and parent report.

The language of the assessment varied, depending upon the needs of the child, the family's level of acculturation, and the language usage patterns demonstrated and/or reported by the family. Because the family's language needs were generally determined at the referral level, assessments were typically conducted in Spanish (N=8). Four children were evaluated in Spanish and English was used as a facilitator, while one child was evaluated in English with Spanish used as a facilitator (Gonzales & Kayser, 1997).

Procedures and Instruments

The investigator's role as participant/observer was integrated into the duties and activities normally conducted as a bilingual SLP. In the majority of cases, the students, families, and colleagues who were the sources for data collection, were typical participants in the special education process; 40 of the participants received speech-languages services with the researcher. This investigator assumed the role of participant/observer in the course of the daily routine of service delivery, assessments, meetings, student/teacher/parent interviews, record reviews, observations, teacher/parent consultations, and report writing.

The initial research strategy involved the collection of authentic assessment data from 65 students who received speech-language services or had been referred for an initial evaluation. Of these 65 students, 40 were enrolled in speech-language services, 12 were school-aged students referred for initial speech-language

assessments, and 13 were preschoolers referred for evaluation by the bilingual preschool assessment team. With the exception of students who had been referred for initial evaluations, authentic assessment procedures were integrated into the delivery of speech-language services (see pages 60-63 for specific examples of authentic assessment techniques). These authentic assessments were completed in both Spanish and English for 14 students who evidenced sufficient cognitive/academic English language proficiency skills (Cummins, 1984) to warrant a possible referral for transitioning into the monolingual program. The collection of assessment data was primarily confined to the Spanish-language for 15 students who were considered to be in the early stages of English language learning. The assessment data were gathered and analyzed on an ongoing basis throughout the research project. Standardized assessments supplemented authentic assessment procedures as needed.

Although the assessment protocol varied depending upon the needs of the individual learner, the authentic assessment procedures included the following:

Communicative/Literacy Portfolios	A collection of student work products that demonstrate a student's progress in understanding and using language for the completion of reading and writing tasks. Portfolio conferences between the student, SLP, parent and/or teacher were conducted in order to discuss progress and plan goals as related to the artifacts collected in the

	portfolio (Kayser, 1998; Kratcoski, 1998; Wiener & Cohen, 1997).
Miscue Analyses	A type of reading assessment that examines a student's use of decoding skills, reading strategies (e.g., contextual cues), and comprehension while reading aloud (O'Malley, 1996).
Running Records	A type of reading assessment that analyzes reading accuracy by examining a student's uncorrected reading errors (e.g., substitutions, omissions, and insertions) while reading aloud (O'Malley, 1996).
Narrative Samples	An assessment procedure that examines a student's comprehension skills and knowledge of story structure by analyzing the ability to retell a story that has been read to or read by the student (O'Malley, 1996).
Interviews	Ethnographic based interviews with the student, teacher, and/or parent in order to identify the components of the curriculum that may be problematic for the student (Nelson, 1992, 1994). Parent and/or teacher and student interviews were also used to gain information about each student's language usage patterns (language proficiency, language dominance, and language preferences) in

	the home, school, and neighborhood settings (Mattes & Omark, 1984).
Language/Literacy Rating Scales/Checklists	Rating scales or checklists that identify target language or literacy behaviors based on teacher report and/or structured observation of the student in the classroom or therapy settings (O'Malley, 1996; Roseberry-McKibbin, 1995; Secord, et al., 1994).
Rubrics	A type of assessment scale used to measure a student's performance. Rubrics are comprised of a fixed scale and a list of attributes that describe criteria at each score point for a specific outcome (O'Malley, 1996).
Language Samples	The systematic collection and analysis of a student's verbal output (Nicolosi, Harryman, & Kresheck, 1978).
Therapy Logs	A brief narrative summary of a student's progress toward the mastery of specific therapeutic goals completed after each therapy session or following consultation with the parent and/or teacher (Flower, 1984).
Acculturation Scales	An instrument used to measure an individual's level of acculturation (see pages 46-49) that may be

> determined through self-reported information about patterns of language preference and use, peer associations, and use of media (Barona & Miller, 1994).

Field notes and reflections by the SLP on the *process* of the data collection were completed several times weekly or as appropriate. Field notes were generated as part of the perusal of cumulative records, formal/informal assessments, and student interviews/observations. Field notes were completed by the investigator following participation in Service Teams, Initial Assessments, and Annual Review Meetings. These notes focused upon the process of assessing students in a more authentic manner and not upon the actual student generated products of these assessments.

Copies of Initial Assessment reports were a component of the data collection process. Written summaries of Annual Review assessment results were collected as well. These reports were gathered in order to analyze and evaluate the process of incorporating authentic assessment into service delivery. Quarterly progress reports were also collected and examined for the application of authentic assessment procedures for the purpose of documenting progress in therapy.

Twenty-three portfolio conferences were held in order to demonstrate students' progress in the therapy and classroom settings and to set goals for future therapeutic services (O'Malley, 1996; Valencia & Afflerbach, 1994). Although these conferences were generally organized to coincide with Annual Review Meetings, four took place in the students' homes. Due to scheduling conflicts and/or difficulties with classroom

coverage, conferences with parents, teachers, and students could not be consistently coordinated. These conferences were attended by a variety of individuals including classroom teachers, English-as-Second-Language (ESL) teachers, teaching assistants, parents, grandparents, one student's counselor, classmates, and of course the students themselves.

Students, parents, and educators were interviewed following their participation in the portfolio conferences. Following review of the students' portfolio, 10 parents and six classroom teachers participated in brief interviews using an ethnographic approach (Bogdan & Biklen, 1992, 1998; Mertens, 1998; Miriam, 1998; Seidman, 1998; Spradley, 1979). These interviews solicited their opinions regarding the use of authentic assessment techniques with preschool and elementary grade students with communicative disorders (see Appendix A for a listing of the questions). A group of 13 students enrolled in second grade and above were asked to reflect on their experiences with communicative/literacy portfolios and conferences. Nine special education personnel with at least one year of experience in the use of authentic assessment techniques were interviewed about the advantages and disadvantages of these alternative assessment procedures. The purpose of these student, parent, and educator interviews was to adapt and/or modify the ways in which authentic assessments could be implemented in speech-language therapy in order to more closely match the student's communicative needs, the instructional activities of the classroom, and the educational goals of the family.

Finally, a group of seven educators including another bilingual SLP met on a monthly basis to explore the use of authentic assessment techniques at the preschool

and elementary school level. The purpose of these meetings/workshops was "to provide educators with a comprehensive, systematic, and integrated framework to assist them in conceptualizing and designing authentic assessments" (Worcester Public Schools Office of Staff and Program Development, 1998, p. 11). Guest speakers discussed the ways in which they had incorporated authentic assessment techniques within their own classrooms. Participants explored ways of adapting the guest speakers' recommendations to their own individualized educational settings. Field notes and reflections were completed by the researcher following each one of these meetings.

The three research questions were answered through the inductive analysis of the data drawn from the field notes, reflections, and interviews. The data were interpreted using the constant comparative method (Bogdan & Biklen, 1992; Glazer & Strauss, 1967; Kayser, Brice, Munoz, Moss, & Davis-McFarland, 1997, Miriam, 1998). Using this approach, the collected data were analyzed for similarities that reflected the general categories related to the research questions. Through on-going analysis between and among the coded categories, this investigator searched for emerging patterns and triangulation of the data from the multi-layered sources of information. This process resulted in the refinement and delimitation of the initial coded categories and the emanation of the key categories that addressed the research questions.

CHAPTER 4
FINDINGS

Restatement of the Problem

As educators face the challenges of the new millennium, the repercussions of the Massachusetts Reform Act of 1993 continue to exert an increasing influence on all aspects of the educational experience of students, parents, and school personnel. Buzz words, such as accountability, high standards, and measurable outcomes, are bandied about by the media, politicians, and the general public; however, it is within educational circles that the consequences of high-stakes are most keenly felt. One of the positive outcomes of both the Educational Reform Act and IDEA-97 is that educators will be held accountable for the educational outcomes of all students, including those with special needs and from culturally and linguistically diverse backgrounds. A major drawback of high-stakes tests is that educators, fearing that their students will be labeled as under-performing, have referred an increasing number of students, especially cultural and linguistic minorities, for special education assessments. Bilingual and ESL educators, whose students must take high-stakes assessments in English if they have lived in the United States for three or more years, are exiting their students as quickly as possible, often before they have sufficient English-language proficiency to pass these tests. There is a mistaken belief that bilingual students with communication disorders are unable to develop proficiency in two languages, so it is best that they be immersed in the English language. These students are often quickly exited into monolingual programs without the requisite

English language skills to do well on these high-stakes tests. Although the consequences of high-stakes assessments will have a major bearing on educational decisions in the near future, the judicious use of authentic assessment procedures such as portfolios hold promise as supplementary devices in interpreting the results of these tests.

This study investigated the use of authentic assessment procedures with bilingual students with communication disorders. The specific purposes of the study were to:

(1) Identify how authentic assessments could be used to determine the language of instruction.

(2) Identify how authentic assessment procedures can be used to document progress in the attainment of speech-language therapy goals as outlined in students' Individual Educational Programs (IEP).

(3) Identify how authentic assessment procedures can be used to supplement standardized assessments for the differentiation of language differences from language disorders.

Development of Authentic Assessment Procedures

The first part of this chapter is a description of the process and products that were developed as part of this investigation into the role of authentic assessment practices in the field of communication disorders. This will be followed by a discussion of how the use of these authentic assessment procedures answered the three

research questions. Final conclusions and recommendations for future research will be addressed in Chapter 5.

Traditionally, speech-language assessment has been an activity that is conducted on an individual basis in a manner disassociated from therapeutic activities. This requires the cancellation of multiple therapy sessions. Although certain aspects of the traditional assessment process must be maintained, especially when working with unfamiliar students, this study represents a work in progress as well as a change in the mind-set of what defines the assessment process.

Throughout the study, the investigator incorporated a teach-evaluate-teach cycle in which assessment became a component of therapeutic activities. Observation of the student's performance during the activity and/or analysis of the artifact generated by the activity guided the subsequent interventions. Assessment was a daily activity that analyzed, synthesized, and documented information from multiple sources and from a wide array of informants, including teachers, parents, and/or students.

Portfolios were developed for all of the students who were enrolled in speech-language services. Work samples generated in therapy were placed in each student's portfolio. A second record-keeping portfolio maintained by the researcher, contained photocopies of the student's artifacts as well as samples from the other authentic sources of information such as running records, miscue analyses, rubrics, checklists, acculturation scales, therapy notes, language samples, interviews, and progress reports.

Communicative/Literacy Portfolios

Communicative/literacy portfolios (CLP) were assembled for all students who received speech and language services with the researcher. These portfolios served three purposes (Wiener & Cohen, 1997). First, the portfolio established baseline information about each student's literacy skills, communicative abilities, and personal interests. Data regarding the student's developmental history, including medical, social, and educational history, were also filed in the record-keeping portfolio.

Second, the CLP provided cumulative information about a student's progress toward the speech and therapy goals as outlined in the IEP. Specific work samples were collected at designated intervals (e.g., fall, winter, and spring). The analysis of specific work samples documented each student's progress in attaining specific speech and language goals as well as his/her development of literacy skills (Wiener & Cohen, 1997).

Finally, the CLP provided a summative assessment of each student's communicative and literacy skills, which were completed prior to the termination of the IEP period. These assessments assisted in determining the need for continued services as well as in developing appropriate educationally relevant therapeutic goals. Summative portfolio assessments also aided in the selection of the suggested language of instruction and/or therapy (Wiener & Cohen, 1997).

Each portfolio reflected the individual communicative and learning needs of the student. The work samples targeted students' communicative as well as their literacy goals. Repeated renditions of specific core activities were used to document progress throughout the school year. One of these core activities was self-portraits

that were completed by the students in the fall, winter, and spring. Students drew pictures of themselves and/or their families as they engaged in a favorite activity and then wrote a brief explanation of the depicted events. Students were tape-recorded as they described their self-portrait.

In the majority of cases, an autobiography replaced the self-portrait as a core activity for students in grades three through six. A process writing approach was employed that involved idea generation and planning, drafting, revising, editing, and publishing. All components of the assignments were filed as evidence of both the process and the product for this activity. Students were tape-recorded as they read their autobiographies and the tapes were presented to their parents, teachers, or peers during the portfolio conferences.

Several students wrote an autobiographical letter to a pen pal in another school. The students followed the process writing approach. Upon completion, the students e-mailed the letters to their pen pals. As these written tasks were a requirement of the district-wide portfolio, these artifacts were also filed in the students' classroom portfolios.

Story maps (pictographs and written output adapted from the Success For All Foundation, 1998), graphic organizers (e.g., word webs adapted from the Success For All Foundation, 1998), story summaries (pictographs and written output adapted from the Success For All Foundation, 1998), and literary responses to stories read aloud in therapy were also filed in the CLP. Students and/or the researcher also transcribed a list of the books read in therapy. Students were invited to include items from their

classroom portfolios in the CLP. All artifacts were dated and briefly described (e.g., the title of the book associated with the artifact).

These artifacts were copied and placed in the researcher's record keeping portfolio. The researcher wrote a brief description of the communicative behaviors demonstrated by the student while completing the activity, which was also placed on the artifact.

The initial portfolios consisted of a file folder decorated by the student. Portfolios were placed in a storage device such as a plastic milk crate or a clear plastic file container that was readily accessible to the students. In later stages of the research study, the portfolios mandated for classroom use by the school district replaced the manila files. These portfolios outlined the obligatory and optional writings samples required of the student. Students and/or educators also listed the books read to or by the student.

A number of artifacts were also placed in the communicative/literacy portfolio. These items included story conference forms, results from running records and/or miscue analyses, language samples, audio samples, checklists, rubrics, results from interviews, acculturation scales, and formal assessments. The process in generating these artifacts is described below.

Story Conferences

Reading conferences documented students' responses to stories that had been read aloud in therapy. This activity was adapted from the listening comprehension component of the Success For All Reading Program (Success For All Foundation,

1998) and attempted to replicate the reciprocal process of the shared reading experience. The purpose of this activity was to improve students' comprehension for reading and listening tasks, increase their knowledge of story structure (e.g., setting, problem, and resolution), and improve expressive language skills through an interactive discussion of the literature. Because assessment was integrated with intervention, stories were chosen based upon the language needs of the students. For example, a story in which the characters engaged in activities was used to reinforce the use of plural verb forms.

The interactive reading process began with a review of the title and author and predictions by the students of what they thought might happen in the story. The events of the story were previewed and then the story was read aloud. Interactive question probes that encouraged predictive and inferential thinking skills were written on "post-its" and placed on the corresponding page of the book. The session ended with a brief review of the story's structure.

During the second session, the story was retold using pictures from the text, sequence cards, or graphic organizers. The content and structure of the story were emphasized. Students indicated what they had liked or disliked about the story and why.

During the third session, an individual story conference was completed. Each student responded to questions about the story's content and structure (the setting, main idea and/or problem of the story, and its ending). In addition, students' understanding of inferential aspects of the story was probed. See Appendix B for an example of a story conference form.

In the initial stages of this assessment technique, students' answers were scored as correct or incorrect, and a qualitative analysis of the students' responses (e.g., the level of cueing needed by the student) was completed. A rubric, adapted from Porch (1967), was developed for the analysis of students' responses to the target questions (see Appendix C). One-on-one conferences were held and the student's responses to the target questions were written verbatim. While the investigator conducted the conference, the other students in the group responded to the story by illustrating and writing what they liked or disliked about the story and why. Following the completion of the conferences, students in the early stages of literacy development dictated their responses to the researcher.

The individual reading conference and the literary response were completed in Spanish at a minimum of three times a year, typically in the fall, winter, and spring. For students who had been exposed to English for two or more years, the interactive story reading was also completed in English. Although students were encouraged to answer the target questions in English, responses in either or both languages were accepted. Students who were recommended for transitioning to the standard program participated in additional reading conferences in English.

An adaptation of the story conference format for students in the fourth grade and above was developed at the end of the research study. This adaptation of longer texts included the administration of a miscue analysis in the first half of the book. At the end of the story retelling, the researcher read aloud the remainder of the story and a second retelling was completed. This analysis of students' listening comprehension

skills was scored in the same manner as the narrative generated by the oral reading task (see page 188).

Early Literacy Assessments

The Observation Survey of Early Literacy Achievement (Clay, 1993) was used to assess students' early literacy skills, regardless of the actual grade placement of the pupil. Results from Clay's Observation Survey were analyzed quantitatively in six areas including the ability to identify letters, read words in isolation, and demonstrate knowledge about print. The Observation Survey was completed in Spanish; however, assessments were also completed in English with six children who had been recommended for transitioning into the monolingual program.

Running Records and Miscue Analyses

Miscue analyses and/or running records were collected on all students enrolled in therapy as well as the majority of students referred for initial speech-language evaluations. The research literature recommends that running records and miscue analyses incorporate unfamiliar materials from the reading text employed in the classroom (Nelson, 1993, 1994, 1998; Weaver, 1994). A number of difficulties arose when attempts were made to carry out this suggestion. Given the constraints imposed by the school setting, it became problematic to secure the written text, copy it, and then conduct the miscue analysis within a timely fashion. This was especially difficult when completing initial evaluations in unfamiliar buildings. In addition, the reliability

and validity of the analysis of the retell for a less than familiar text was an area of concern. As a result of these difficulties, the researcher established a portable reading assessment kit.

Books were collected in Spanish and English that served as benchmarks for the assessment process. Because multiple factors (e.g., a student's background knowledge) can influence the readability of a given text for a particular student, the reading level of the benchmarked books was established in a number of ways. One way of establishing the initial reading levels of the benchmarked books was through the reading levels that were assigned to specific texts by reading researchers (Beaver, 1997; Fountas & Pinnell, 1996; Routman, 1994). The reading level of these texts was further verified through the consensus of three to five bilingual educators. Teachers were selected who had taught within one year of the book's reading level for at least two years (e.g., second grade teachers judged books from first through third grades). Each teacher was requested to identify individually the reading level of the text in question. If three teachers agreed, the book was considered to fall at that grade level. If disagreement was noted, two additional teachers were consulted and the average of their responses became the book's reading level.

The leveling system designated by Fountas and Pinnell (1996) was eventually adopted. This system labels the reading levels of books using an "A- R" designation system for kindergarten through fourth grade. Because the authors' reading levels did not extend beyond fourth grade, a fifth grade reading level was added and designated as "S". According to Fountas and Pinnel (1996), books designated as "A" or "B" are leveled at a kindergarten/grade one reading level. These books focus on a single idea

or have a simple story line. There is a direct relationship between the text and the pictures. The print appears at the same place on every page and there is sufficient space between words to allow students to easily point to the words while reading. In contrast, Level "M" books are at the second grade level. These books have a large amount of text per page and smaller print. The language structure is more complex and the vocabulary is sophisticated. The themes and concepts presented in these books are more abstract and the subtleties of these texts require background knowledge.

Two to four books were collected per grade level in Spanish and English. The benchmarked book and a transcript of the text were placed in a zip-lock type plastic bag. These books were then stored in a 21-pocket portable filing system that could be easily transported from school to school. A listing of the benchmarked books can be found in Appendix D.

Texts from 20-words to 750-words in length were used to administer the oral reading samples. Whenever possible, a book was chosen that could be read in one setting, thus yielding an oral reading sample and retell on a complete text. In cases where this was not possible (e.g., shorter, well-written, and unfamiliar texts at the appropriate reading level could not be located), the reading of the text ended at natural breaks in the story line. Books were chosen that were unfamiliar as well as challenging to the reader. Texts were selected that were judged to be slightly above the students' instructional reading level based upon teacher report and/or an examination of classroom artifacts (e.g., reading book or writing samples). The appropriateness of a given book was further verified by the behaviors demonstrated

while reading aloud the first few pages and/or paragraphs of the text. If the text seemed to be too easy or difficult, other books were selected until the appropriate level was ascertained. Running records and/or miscue analyses were completed in Spanish two to three times during the school year, typically in the fall, winter, and spring. At least two miscue analyses or running records were also conducted in English for students recommended for transitioning into the monolingual program.

A typed copy of the text to be read by the student was prepared by the researcher for the transcription of the reader's miscues. This copy was double-spaced to allow sufficient space for the marking of all miscues. The miscue analysis or running record was transcribed during the oral reading; however, the session was taped for follow-up analysis.

An adaptation of the transcription system developed by Goodman, Watson, and Burke (1987) was utilized to mark the student's miscues (see Appendix E for an explanation of these conventions). The questions developed by Weaver (1994) were used to analyze how the miscues related to the language that preceded and followed the miscue. Weaver's analysis system was selected because it focused on the connection between reading miscues, oral language skills, and comprehension. The Language and Reading Observation Guide was adapted from the work of several authors (Beaver, 1997; Burke, 1980; Clay, 1993; Weaver, 1980, 1994) for the recording and qualitative evaluation of the behaviors observed during the oral reading and retelling (see Appendix F). The researcher summarized and recorded the observed behaviors and strategies used by the reader on the Language and Reading Observation Guide.

Following the oral reading of the text, students were instructed to retell the story in their own words. A written summary of the story (characters, setting, major events, supporting details, and ending) was prepared that served as a guide for the retelling. In order to recall specific details of the story, the researcher reviewed the summary prior to initiating the miscue analysis. The highlights of the retelling were transcribed in the appropriate place on the Language and Reading Observation Guide. The content and organization of the narrative, the reader's literal and inferential understanding of the story; vocabulary and morphosyntactic skills were analyzed via a rubric adapted from Weaver (1980, 1994) (see page 188). Question probes, based upon what the reader had said about the story during the unaided retelling, were used to evaluate the reader's understanding of items not mentioned or to stimulate elaborated responses for ambiguous statements provided by the student. During this portion of the assessment, caution was exercised to avoid the provision of information not stated by the student in the retelling.

In cases where students' comprehension evidenced an active understanding of the text during the oral reading component but were unable to provide a satisfactory retelling of the story, a silent reading of the text was completed. The student's level of understanding of the text following the silent reading was assessed. The student's retelling was qualitatively judged using an adaptation of criteria developed by Beaver (1997). The follow-up questions were given at the end of all story retellings, while the reading interview was usually completed as part of initial or annual review assessments.

Language and Narrative Samples

Language samples were collected from all students during initial speech-language assessments. These samples were collected within a variety of communicative contexts (e.g., the classroom versus the lunchroom) and/or with a variety of interactors (e.g., peers, siblings, and parents). For school-aged children, language samples were gathered during an array of activities including conversational probes, narrative tasks, or structured assessment procedures. Language samples for preschoolers were usually collected during play activities using a variety of props (e.g., food items in the dramatic play area, toy cars, or farm animals). The child's interests or preferences established the topics of conversation as well as the language(s) utilized during these conversations. Parents usually assisted in the elicitation of the language sample.

During initial speech-language evaluations, the language samples of younger children (preschool and kindergarten levels) were analyzed using mean length of response (MLR). In addition, the quality and complexity of the output as well as the variety of pragmatic language functions (e.g., protests, requests actions, informs) expressed by the student were analyzed. For older students, terminal units (TU) were used to determine the complexity of the student's language (Hunt, 1965). The language samples were analyzed quantitatively (Restrepo, 1998) and qualitatively based upon developmental data for Spanish-speaking students from monolingual and bilingual environments (Anderson, 1995; Kayser, 1998; Merino, 1992).

Language samples were also collected as part of the annual review process. In anticipation of this process, the samples were gathered on an ongoing basis throughout

the year with an array of communicative partners (e.g., classmates, teachers, or friends) and within a variety of contexts (e.g., in the therapy room, classroom, or cafeteria). Given the difficulty of collecting accurate samples while engaged in the dynamics of group therapy, language samples were collected in a cyclical manner. Specifically, a particular student was targeted on a revolving basis and samples of three-to-five utterances were transcribed and dated. This allowed the researcher to collect a representative sample of 50 to 100 utterances for each student over the course of several months. The language samples were transcribed on a 6-inch by 4-inch "post-it" note secured to a small clipboard and later filed in the record-keeping portfolio.

In the majority of the cases, language samples were collected primarily in Spanish. For comparative purposes, less representative language samples (approximately 25-75 utterances) were gathered and analyzed in English. Complete language samples were collected in both languages for students recommended for transitioning into the standard program.

Students' narrative language skills were also assessed following the viewing of a wordless book. Students first looked at a wordless picture book, such as *One Frog Too Many* (Mayer, 1975), in its entirety and then returned to the beginning of the book and told the story page by page. The student's narrative was taped and analyzed qualitatively using an adaptation of a rating scale developed for writing samples (Singer & Bashir, 1998; Worcester Public Schools, 1995). The narrative was further analyzed in the areas of content and structure using developmental data cited by Westby, Van Dongen, & Maggart (1989). These alternative assessment devices

can be located in Appendix G.

The ability to generate expository narrative skills was evaluated for students in the third grade and above. Although the selected activities varied depending upon the specific concerns expressed by teachers, parents, or the student, targeted genres included procedural (explaining how to do or make something), problem-solving (providing a solution to a hypothetical problem), or cause and effect (offering reasons of why something may have happened). These expository narratives were also evaluated using the rating scale (adapted from Singer & Bashir, 1998; Worcester Public Schools, 1995) and checklist (adapted from Westby, Van Dongen, & Maggart, 1989) located in Appendix H.

Researchers have found the use of videotapes as a diagnostically promising means for narrative elicitation (Gutierrez-Clellen & Heinrichs-Ramos, 1993). This idea was adapted using CD-ROMs (Mayer, 1995) that are available in both English and Spanish. For younger students (preschool through second grade levels), the narrative was collected as the researcher and student watched the CD-ROM. Older students, (third grade and above) were asked to view the CD-ROM alone and retell the story to the researcher who pretended that she was unfamiliar with it. The use of CD-ROMs yielded a high level of verbal output from students; however, further investigation of this medium as a language elicitation tool is warranted.

Audio Samples

Audio language samples were another component of the portfolio. These samples typically contained verbal descriptions of students' self-portraits, oral reading

samples, and narrative samples. Preschoolers were taped during play activities in the classroom or therapy setting or while describing work samples. Students with significant speech impairments were tape-recorded as they engaged in an easily replicated task (e.g., counting to ten, repeating a sentence that contained targeted speech sounds). Students with fluency disorders were recorded during reading and speaking activities within the therapy setting as well as under more stressful conditions such as making a phone call or while talking to an unknown adult.

Therapy Notes

Therapy notes were transcribed for each student after each therapy session. During the therapy session brief notes that described the student's overall performance on the targeted tasks were jotted down on a "post-it" note. Behaviors related to the student's progress toward therapy goals were also recorded. At the end of the school day a longer note based upon the "post-it" was transcribed for each student. These notes also contained information reported by the teacher about the student's progress in the classroom or other factors related to service delivery (e.g., canceled sessions secondary to meetings, absences, holidays). During the school year, these notes were attached to a copy of each student's therapy plan that served as a reminder of his/her specific goals. During the school year these notes were placed in a color-coded file, one for each of the four schools serviced. Therapy notes were filed in each student's record-keeping portfolio at the end of the academic year.

Interviews

Nelson (1994, 1998) proposed that "zones of significance " or problematic contexts within the curriculum for the individual student, be identified through the use of ethnographic interviews with the student, teacher, and/or parent. Based upon this recommendation, interviews were completed with 87% of the students (N=45) enrolled in school-based programs. Interviews were conducted with teachers for all of the students as well as 85% of the parents (N=38). Of the 23 students referred for initial speech-language assessments, 48% of the students (N=11) and 83% (N=19) of their parents were also interviewed. Interviews were completed with the parents of all of the students seen through the preschool assessment arenas. Students at the preschool developmental level were not interviewed. Interviews were not completed in cases where parents could not be contacted or when they were unable to attend the Evaluation Team Meeting.

Because the family's level of acculturation was completed as part of the interview, results from this assessment guided the language that was used for the interview. In cases where this information was unavailable prior to the interview, the investigator requested that the interviewee select the language in which he/she felt most comfortable communicating.

Examples of interview questions for students, teachers, and parents can be found in Appendix I. These interview questions, adapted from Nelson (1994), served as a guide for possible topics to be explored during the ethnographic interview. The purpose of the interviews was to examine the interaction of the student within the context of the learning environment from the point of view of the student. Topics

discussed were not limited to the parameters typically associated with speech and language skills, but attempted to explore all aspects that potentially impacted the student's academic performance (Nelson, 1994). The majority of these interviews were conducted as face-to-face conversational interactions. Although the interview questions provided a framework for possible areas of discussion, the interests and concerns of the interviewee guided the course of the interview.

An interview question format was adapted from Burke (1980) for students with documented comprehension and memory deficits, as well as word retrieval difficulties. The purpose of these interviews was to gain insight about the student's use of specific comprehension, memory (e.g., visual imagery, "chunking" information into smaller more easily recalled units), or word retrieval (e.g., describing a word that cannot be immediately retrieved via its attributes) strategies. This format was also adapted for one student who presented with a language-based fluency disorder (cluttering). As one of the characteristics of this disorder is a decreased level of awareness of the communicative difficulty, this interview provided evidence of the student's level of understanding of his/her ability to cope with the communicative challenges of the home and school settings.

Language and Literacy Rating Scales and Checklists

Several checklists and rating scales were used as measures of students' communicative skills as well as their ability to access the classroom curriculum. The Curriculum-Based Reading and Language Arts Inventory (CRLAI) (Worcester Public Schools, 1999) was developed from the 28 Learning Standards of the Massachusetts

Curriculum Frameworks (for example, learning standard number one is that students will follow agreed upon rules for classroom discussion and carry out assigned roles in self-run small group discussions). The inventory is appropriate for use with students in preschool up to the 12th grade and requires that the teacher indicate the presence or absence of the targeted behaviors. The CRLAI served as a point of reference when interviewing teachers about students' ability to use language to effectively access the curriculum.

The Classroom Communication and Learning Checklist (CCLC) (Wiig & Secord, 1994) was used when students demonstrated both communicative and academic concerns. This checklist, developed by the authors in collaboration with elementary classroom teachers, identifies 66 target behaviors related to language and learning difficulties. The frequency of occurrence of these behaviors (never, rarely, some of the time, and most of the time) and the level of the educator's concern about these behaviors (none, only a little, some concern, and very concerned) are rated. In cases where the referring teacher was unable to specifically delineate areas of concern, this checklist was utilized to establish the "zones of significance" (Nelson, 1994; 1998) so that the assessment protocols could be individualized to the student's curriculum-based language needs. The CCLC also served as a structured interviewing device in cases where a comprehensive evaluation could not be completed prior to the Evaluation Team Meeting.

An abbreviated version of the Bilingual Classroom Communication Profile (Roseberry-McKibbin, 1995) was developed. This checklist examines a variety of cultural and linguistic factors and their influence upon a student's classroom

communication and learning behaviors. This checklist was typically used with bilingual students enrolled in the monolingual program, especially in educational environments that were not typically serviced by the researcher.

The Observation of Reading Behaviors was adapted from Fountas and Pinnell (1996). This rating scale was based upon the authors' description of specific behaviors that the observer should notice and support at different stages in reading development. For example in Level F, the student should be encouraged to search visual information to figure out novel words when reading. This rating scale was used in conjunction with a running record or miscue analysis. The targeted behaviors were rated as beginning, developing, or well developed. A copy of this rating scale is included in Appendix J.

Given the relationship between language and reading disorders, the Early Identification of Language-Based Reading Disorders: A Checklist (Catts, 1997) was used to identify children in kindergarten and first grade who might be at-risk for reading difficulties. This checklist addresses those areas associated with language-based reading disorders such as word retrieval, auditory memory and speech production difficulties. This checklist was completed with input from kindergarten and first grade teachers.

Rubrics

The Student Writing Rubrics (Goddard School of Science and Technology Faculty, 1996) was developed by a group of monolingual and bilingual educators at an elementary school (preschool through sixth grade) in the district in which the research

study took place. This rubric served as the basis for a narrative summary of students' writing skills as demonstrated in their work samples (e.g., self portraits, autobiographies, letters, literary responses). Comments about both the process and product of the written samples were transcribed on a "post-it" note that was affixed to the researcher's copy of the artifact. For example, early writing behaviors for a kindergarten student included using pictures in a detailed manner; writing name and favorite words; and using words to convey a message.

The Massachusetts English Language Assessment-Oral (MELA-O) (1994) is used by ESL teachers to qualitatively describe students' oral language skills. Students' comprehension and production (fluency, vocabulary, pronunciation, and grammar) skills are rated for six levels of proficiency. This rubric was adapted for students who were exposed to Spanish in the home setting, were not enrolled in the bilingual program, and were experiencing language loss. It was used to qualitatively describe students' Spanish language skills as observed during the evaluation. For example, a student at Level 3 would demonstrate an understanding of most interpersonal interactions and be able to use basic grammatical forms. The observations from the evaluation were corroborated with information obtained through interviews (e.g., student, parent, former classroom, and ESL teachers), record reviews, and the completion of the acculturation scale.

For students who had been recommended for transitioning into the monolingual program, samples from the their classroom portfolios were analyzed using the Holistic Scoring Rubric for Writing Assessment with English Language Learning Students (O'Malley, 1996). This rubric, developed by ESL teachers from

the Prince William County Public Schools in Virginia, was designed for use with students in the process of learning English. It rates a writing sample in five dimensions (meaning, organization, use of transitions, vocabulary, and grammatical/mechanical usage) and for six levels of development from emergent to proficient (O'Malley, 1996).

Acculturation and Social and Educational History

Information about each student's level of acculturation was included in the portfolio. For school-aged students, this process typically began with a record review that focused on the student's and parents' birthplace, the family's socioeconomic level (e.g., parental employment), and the student's educational history (e.g., length of time enrolled in the bilingual program). The Language Preference and Home Language Survey Form (Worcester Public Schools, n.d.), a questionnaire that is completed by parents regarding the language usage patterns of the home, was also examined. Records related to bilingual (comments by the teacher regarding the student's English-language proficiency) and/or ESL (e.g., results from the MELA-O) programs were also perused.

Kindergarten or first grade students who had been exposed to a limited amount of English prior to their placement in the bilingual program or students recently arrived from a Spanish-speaking country were informally screened. The acculturation screening consisted of interviews with the student, teacher, and/or parents about the child's pattern of language use in the home, school, and neighborhood settings. The results of the screening were supplemented with a review of the student's cumulative

record. As the majority of these students presented with a low level of acculturation, further testing was not typically warranted.

Students who were enrolled in second through sixth grade were administered the Short Acculturation Scale for Hispanic Youth (SASH-Y) (Barona & Miller, 1994). Students, teachers, and/or parents were interviewed about the pupil's pattern of language use, proficiency, and preference for communicative and academic purposes in the home, school, and neighborhood settings. The SASH-Y was administered as part of an initial speech and language evaluation as well as the annual review assessment process.

Students in kindergarten or first grade who had been exposed to English for an extended period of time (e.g., both languages were spoken in the home) were given an adapted version of the SASH-Y. The students first drew pictures of people associated with the language options on a horizontal continuum (A= a person familiar to the student who speaks exclusively Spanish, C= a person familiar to the student who speaks both languages, and E= a person familiar to the student who speaks English), and then were asked to compare their language use to that of the known person. When students reported that they used both languages, they were asked to further qualify their responses. Other students who evidenced difficulty with the multiple-choice format were given an adapted version in which the options were decreased to three (A= Spanish, C= Both, and E= English) and their answers were scored accordingly (i.e., A=1, C=3, and E=5).

The wording of the questions was changed to fit the needs of younger students. For example, in the first question students were asked if they understand and speak the

languages instead of read and speak the languages. For question 11, children were asked if they liked to play with Puerto Rican children (or children of their own ethnic background), non-Puerto Rican children, or both. Results were considered to provide an estimation of the child's level of acculturation. The results from the screening were corroborated with reports from family members and/or teachers as well as from case history information (e.g., place of birth or parents' places of birth).

Because each family's language needs were determined at the referral level, children seen through the preschool arenas were typically from homes in which predominantly Spanish was spoken. The child's level of acculturation was established through observation and parent interview (e.g., the patterns of language exposure and use in the home, neighborhood, and/or preschool settings, the place(s) of birth of family members). In cases where children had been exposed to both languages, parents were asked to identify the language(s) the child understood or spoke most proficiently (e.g., when you really want your child to understand something which language(s) do you use? Which language(s) does your child speak when he/she really wants you to understand something?). Parents were also asked to describe the way in which a second language may have been introduced (e.g., through a day care setting) and the effect on the child's first language.

Social history information was compiled via record reviews as well interviews with the student, parents, and/or former teachers. When reviewing the student's cumulative records, the researcher attempted to maintain an ethnographic perspective regarding the reported results as well as the influence of possible cultural and linguistic variables on the reporter's perspective. Whenever possible, information

contained in student records was triangulated with information provided by the parent or student. Because a student's level of acculturation was determined on a yearly basis, this form, which served as a permanent record of the student's social, educational, and cultural history, was stapled to the inside cover of the record-keeping portfolio (see Appendix K).

Formal Assessments

The results of formal assessments were filed in the record-keeping portfolio. Although the purpose of the study was to investigate the use of authentic assessment procedures, formal testing procedures were not abandoned. Formal testing procedures were used to complement data obtained from authentic assessment results, so that each source of information qualified and clarified the results obtained from the other. In many cases, the quantifiable data provided a baseline for interpreting results from the more authentic information sources such as the acculturation scale, parent/student interviews, classroom observations, and record reviews.

Portfolio Conferences

The culmination of the authentic assessment process was the portfolio conference in which the student presented carefully selected artifacts to the parents, teacher, and/or the SLP. This process required careful planning and preparation on the part of both the student and the investigator. Based upon the advice of Wiener and Cohen (1997), the researcher incorporated four critical elements into the conduction of

portfolio conferences. These elements included collecting, collaborating, consulting, and communicating.

Collecting refers to the selection of materials that will be placed in the portfolio and thus presented during the portfolio conference (Wiener & Cohen, 1997). According to De Fina (1992) the differences between a portfolio and a folder lies in the student's process of reflection. A folder is merely a collection of work samples, while a portfolio contains carefully and reflectively selected pieces. Because students enrolled in speech-language therapy were seen for a small fraction of the school day (no more than two times weekly for 30-minute sessions), the work samples filed in the CLP were limited in number. The majority of the work samples completed in therapy remained in the CLP; however, students selected those pieces that they wished to present during the portfolio conference.

Students' participation in their own learning is an essential component of authentic assessment; therefore self-assessment and reflection were woven into all aspects of the portfolio process. Because self-reflection is a metacognitive skill that requires students to think about their own learning, this component of the portfolio process was especially difficult for the majority of the students. Based upon interviews with students and teachers, it was evident that many of the students did not have experience reflecting about the process of their own learning. Initial efforts required a high degree of adaptation and scaffolding by the investigator, with slow improvement noted as the year unfolded.

In an effort to improve students' self-reflection skills, an exercise was employed based on the Know-Want to Know-Learned (K-W-L) model. Students in

second grade and beyond identified their present abilities in the areas of speaking, reading, and writing in a sentence completion task. They then wrote goals that stated what they wished to achieve and their success in reaching their goals was reviewed at the end of the year.

Students periodically examined their work samples and reflected on the learning process associated with the artifact. Younger students chose one of three "faces" (happy, "OK", or sad) to indicate their feelings of satisfaction regarding their level of effort or the degree of learning associated with the completion of an artifact. Older students evaluated their own work as well as that of their peers. In this collaborative self-assessment, students stated one thing that they liked about the work sample and one thing that they thought the person might have learned by doing it. At the end of the marking periods, students perused their portfolios and identified areas of growth.

Collaborating with parents, students, and teachers was a key component in the development of the portfolio (Wiener & Cohen, 1997). This required a collaborative approach to learning outcomes as students, parents, and/or teachers identified targets to be addressed in therapy. The culmination of the collaborative assessment approach was the portfolio conference that included parents and/or teachers, students, and the researcher.

Prior to the portfolio conference, each student selected specific work samples to present to their audience. They also prepared a verbal or written explanation of why the particular artifact was chosen. This process was completed in a collaborative

manner in which peers assisted one another with the self-assessment and reflection processes.

Consulting with parents, students, and teachers in order to gain their perspective on the child's communication and learning needs was a key element in the authentic assessment process and therefore the portfolio conference (Wiener & Cohen, 1997). In contrast to the deficit model that only looks at what the student cannot do, consultation provided information about the student's growth in the home and school settings and focused on what the student needed to do to function successfully within those contexts. The incorporation of consultation into the portfolio process provided a balanced view of students' disabilities within the context of their abilities.

Communicating with their parents and/or teachers about their learning as represented by the artifacts in their portfolios was the essence of the portfolio conference for the students (Wiener & Cohen, 1997). The conferences were planned to coincide with the annual review meeting. During the annual review meeting, the student's progress toward the acquisition of communicative targets was discussed and new therapeutic goals were developed. Traditionally the annual review meeting consisted of the presentation of test scores that focused on the pupil's deficits. Although the pattern of the student's strengths and areas of need were discussed during the portfolio conference, the focus was on what they had accomplished and how these skills had related to the classroom curriculum.

With the exception of those 14 years of age and older, students are typically excluded from the annual review meeting. In portfolio conferences, students were active participants in the process as they displayed their work samples and explained

how the artifacts demonstrated their achievements within the therapy and classroom settings. Rather than scores on norm-referenced tests, parents were able to see and hear what their child had accomplished over the course of one year. For example, parents listened to a taped oral reading sample, viewed a picture that summarized the major aspects of the story, and listened to the student's explanation of why the work was chosen for presentation.

Portfolio conferences were conducted with 58% (N= 23) of the students enrolled in speech-language services. Three conferences were attended by the student, parents, classroom teacher, researcher, and, on one occasion, a student's counselor. On three occasions parents indicated that they were unable to attend so the conference was held with the teacher, child, and SLP. In instances when classroom coverage could not be secured, the students presented their portfolios to the parent(s) and the researcher. Two students presented their portfolios to their ESL teacher, the classroom teaching assistant and nine of their peers who proved to be especially supportive of their efforts. Following this presentation, the classmates identified the positive aspects of the presentation as well as how the work samples reflected the students' learning outcomes.

At the end of the school year, students who would continue speech-language therapy with the researcher selected two or three pieces of their work for placement in their new portfolio. These students completed a form or provided dictated responses as to why these pieces were chosen for inclusion in their new portfolio. Two copies of these work samples were made. One was placed in his/her classroom portfolio and one copy was sent home with the student. In cases where students continued with

therapy, but not with the investigator (they were transitioned into the monolingual program), a copy of the portfolio was forwarded to the receiving SLP with an explanation of the contents.

Parents were interviewed about their perceptions of the portfolios and portfolio conferences (see Appendix A for a listing of the questions). Of the ten parents interviewed, nine cited the taped reading sample/retell as the best part of the conference because it most clearly demonstrated academic growth (reading fluency and comprehension). Several parents also indicated that they liked the taped language sample as it evidenced their child's improved speaking abilities. According to eight parents, the work samples showed their child's improvements in writing and drawing skills. Three parents reported an increased level of interest in therapy by their children as a result of the communicative/literacy portfolios.

All of the interviewed parents who had previously attended Initial Evaluation or Annual Review Meetings indicated that the portfolio conference was much more meaningful than the other special education meetings. Parents described the portfolio conference as "logical", "more understandable", and "impressive". In one case a parent who had attended approximately five special education meetings indicated that she was unable to recollect specific details from these meetings. Following reference to her child's participation in portfolio conferences in another school, she immediately recalled these conferences, thus suggesting that much of the information provided during special education meetings may not be presented to parents in a comprehensible manner.

In the majority of the cases, parents declined to offer suggestions for the improvement of the portfolios and portfolio conferences. One mother recommended that the researcher document students' growth by taping a student during an initial reading of a passage and then asking the child to reread the same passage during the portfolio conference.

Students were also interviewed about their experiences with portfolios and portfolio conferences (Interview questions can be found in Appendix A). All 13 of the students expressed positive views of both. Ten of the students described the audio sample as the best part of the portfolio because it demonstrated their progress in reading and "talking" and pleased their parents and/or teachers. According to one student who was taped as she worked with a preschooler, the taped language sample was the best because it provided evidence of what she had taught the preschooler and demonstrated that she could "talk slowly and read perfect(ly)".

The majority of students also indicated that they liked their drawings (e.g., literary responses and story summaries) and their writing samples. In one case a student stated that he liked to draw and write things about books that he had read. One student who had not been enrolled previously in a bilingual program reported that she was proudest of her autobiography written in Spanish.

Ten of the interviewed students had participated in portfolio conferences. When asked to describe how they had felt during the conference, most reported feeling nervous, scared, or embarrassed, as well as proud and happy. One student reported that he was "so happy" that he could show his parents how well he reads and talks. He added that because his parents meet with the teachers and he is never allowed to attend

the meetings, he thought it was great that he "could be there, too". The two students who had read their autobiographies to their peers described feelings of pride, especially due to the high level of positive comments offered by their peers under the direction of the ESL teacher.

Although three individuals thought the portfolios were "perfect" and "didn't need to be improved", the interviewed students provided a number of suggestions. Nine students suggested that the portfolios should include more writing samples, while ten thought that more drawings should be included. A number of the students recommended the addition of supplies (e.g., pencils, paint, magic markers, tapes, books). One student suggested that the portfolios be done on a daily basis. His classmate advised the completion of portfolios in both languages, for those who "don't know Spanish" and that they should be "shown to friends". All of the students reported that portfolios should continue to be used in speech-language therapy.

Five of the interviewed teachers were participants in portfolio conferences with the researcher. When asked their viewpoints about the CLP and portfolio conference, all of the teachers indicated that they appreciated the relationship of the CLP to the classroom curriculum. A teacher trained in the use of running records reported that sharing information about her students' reading skills with someone with a different type of training from her own was informative and gave her a new perspective on the abilities of her students. Another teacher indicated that she liked the way that the portfolios portrayed the "whole child" and didn't focus only on the students' deficits. Two teachers commented that they liked the emphasis that portfolios place on

students' strengths, an aspect that they thought was important for the development of self-esteem.

In the following sections, the results from interviews with 15 educators from special and regular education programs will be applied to the three research questions. This will be followed by the researcher's reflections and findings on the applicability of authentic assessments procedures to differentiate a language disorder from a difference, to determine the language of instruction, and to document progress in the attainment of therapeutic goals. A summary of the findings and recommendations for future research will be addressed in Chapter 5.

Differentiation of Language Differences and Language Disorders

Analysis of the data from the teacher interviews revealed the emergence of four major categories that addressed the research question of how authentic and/or portfolio assessments could be used to differentiate language differences from language disorders. According to the informants, the ideal authentic assessment model was curriculum-based, multidimensional, and student centered. Although the informants strongly supported the implementation of the authentic assessment model, the section ends with a description of the difficulties associated with this model as reported by the interviewees.

The majority of the interviewed teachers emphasized the value of SLPs' incorporation of students' ability to access the classroom curriculum into their assessments. One educator suggested that SLPs examine the language underpinnings of the curriculum in order to determine if students had the requisite skills to succeed in

the classroom. Several suggested that SLPs work with students within the classroom to increase their understanding of the curricular expectations and to see first hand how students' communicative skills compared to those of their peers. Others advised that SLPs observe how students functioned in the classroom to get an idea of the impact of their communicative difficulties on their classroom performance.

One SLP described authentic assessments as a functional overview of students' educational performance that provided baseline information for the comparison of students' performance on standardized assessments with their performance as demonstrated in the classroom. Several SLPs reported that the analysis of the linguistic underpinnings of students' work samples and the observation of their communicative abilities in the classroom provided the information necessary to determine how students functioned in the classroom. According to one SLP, running records provide more information about students' ability to understand and use language than standardized assessments do.

The importance of obtaining data about the student's communicative abilities from multiple sources of information was cited by many of the informants. A preschool teacher urged SLPs to devise ways of systematically collecting information within naturalistic contexts, "because if you just have one piece of information then you don't have an accurate picture of the child's communicative abilities". She suggested that SLPs examine work samples, complete developmental checklists, and observe children multiple times to get a holistic picture of how they interact with different aspects of the curriculum. Several educators stressed the value of

collaborative approaches that incorporate interviews with teachers and parents in order to understand the needs of the whole child.

Many of the informants stressed the merit of observing students as they interacted with various components of the classroom curriculum. A preschool teacher urged SLPs to observe children in different settings and at different times of the day in order to understand their communicative abilities. Others spoke of the importance of observing children outside of the classroom, such as during recess or lunchtime. A SLP reported that the observation of some of her students in the home setting provided her with the perspective of how they communicated in the context of known routines and with familiar interactors.

The importance of interviews with parents and teachers was also cited. The majority of the informants stressed the value of parent interviews as a means for obtaining information about children's developmental histories and for identifying families' concerns about students' communicative abilities. Several SLPs reported that they base their assessment protocols on concerns expressed by teachers, while others indicated that parents serve as resources for goal setting and program planning.

Many informants recommended language samples as an accurate means for assessing students' language skills. An ESL teacher advised that this process should be conducted in an informal and non-threatening way and that the evaluator should follow students' lead by discussing topics that are important to them. A preschool teacher suggested that language samples be collected during art activities, circle time, or lunchtime in order "to get a better idea of how students really communicate". A

number of informants indicated that the incorporation of technology such as video or audio recording would facilitate the process of language sampling.

A SLP reported that she assessed students' narrative language skills using wordless books, story retells, or descriptions of personal events. She analyzed the narrative's cohesion, coherence, and complexity. Although she examined the impact that the grammatical errors exerted on the comprehensibility of the narrative, she was usually more concerned with its content. She reported that when she completed authentic assessments such as story retells, her students didn't really seem to be aware of the fact that they are being tested so there was much less pressure on them than with norm-referenced tests.

Several educators extolled the benefits of dynamic assessment techniques over static testing procedures for the differentiation of language differences from language disorders. One educator said that she used a "testing to the limits" approach that provided qualitative information about students' potential. She described this information as more valuable than scores or figuring out if students "passed" the test. Other educators stressed the merit of teaching students a new task in order to determine the level of effort needed to learn the task. In order to distinguish language disorders from differences, a SLP suggested that ideally services should be provided on a trial basis to determine students' rate of progress and potential for change.

Another SLP praised the way in which dynamic assessments differentiated language disorders from differences by demonstrating students' potential for change. Although characterized as time consuming, she indicated that dynamic assessments were especially useful with children who had done poorly on standardized tests and

were having difficulty in the classroom, because they showed if students had the potential to learn new information at an average rate and were therefore not disabled.

One SLP reported that she regularly had adapted the rubric from the Porch Index of Communicative Ability (Porch 1967) to analyze the amount and type of cueing needed by students on standardized tests (see Appendix B for an abbreviated version of this rubric). This informant indicated that she first gave the norm-referenced test as directed in the testing manual and then re-administered the test while providing scaffolding. This adaptation provided her with information about students' potential and identified the best ways to facilitate students' learning in the classroom.

The importance of investigating students' language background as part of the process of distinguishing language differences from language disorders was cited by many of the informants. It was advised that a record review (e.g., the Language Preference and Home Language Survey Form, Worcester Public Schools, n.d.) be completed in conjunction with an interview with family members about the language usage patterns of the home. Additional suggestions for topics of discussion during the interview included the manner in which the student was first exposed to the second language (e.g., through day care) and the parents' impressions of the child's proficiency in the first, second, or both languages. Others suggested interviews with students concerning their views of their language proficiency, use, and preference as well as their level of comfort when communicating in the first, second, or both languages. A SLP argued that examiners must be cognizant of the difference between proficiency and preference as well as the impact that classroom expectations exert on students' language behaviors.

Other sources of information for the completion of a multidimensional assessment included a review of students' educational history and the analysis of cultural factors. Many of the interviewed educators emphasized the importance of investigating students' academic history and classroom experiences via record reviews, parent interviews, consultations with previous teachers, and student interviews. In cases where students were experiencing academic difficulties, several informants stressed the need to determine their exposure to the targeted tasks.

Although a limited number of the informants actually used the term "acculturation", many referred to the importance of speculating on the influence of culture on students' performance on tests or academic tasks. A bilingual teacher reported that prior to initiating a referral to special education, she investigates the student's background history (e.g., the student's place of birth and the educational level of the parents) in an attempt to determine the possible reasons for the student's academic difficulties. Several SLPs stressed the importance of considering cultural differences when evaluating students from culturally and linguistically diverse backgrounds. Two suggested that evaluators consult with students' parents and/or cultural informants prior to attempting to assess students from a culturally and linguistic diverse backgrounds. In addition to cultural and language differences, a SLP argued that socioeconomic factors must also be considered during speech and language evaluations.

Despite the diverse work experiences of the interviewed educators, the overwhelming majority argued that both standardized and authentic assessments were needed when making educational decisions, especially for students with special needs

and those from culturally and linguistically diverse backgrounds. According to one preschool teacher both standardized and authentic assessments should be completed, because the information from each source enhanced and supported the information from the other source. An inclusionary teacher argued that more accurate assessment results were obtained when formal tests were supplemented with alternative assessments such as portfolios and rubrics, especially when making high-stake educational decisions

The value of pairing standardized and authentic assessments was echoed by assessment personnel. One SLP argued that more subtle learning needs might not be identified with the exclusive use of authentic assessment procedures. Another SLP supported the need for some sort of standard, especially when trying to determine students' eligibility for service. A school psychologist reported that although she was required to use norm-referenced tests, she supplemented this information with data from other sources such as "testing to the limits" so that she has quantitative and qualitative information about students' learning styles, strengths, and weaknesses. A SLP warned that the exclusive reliance on authentic measures could be "very dangerous", especially when assessment personnel were uninformed about "what is normal".

The informants praised the student-centered focus of authentic assessments such as portfolios. A preschool teacher, experienced in the use of portfolios, stressed that students gained a sense of ownership of their own learning when they participated in the selection of artifacts for their portfolios. A SLP suggested that students should be informed of the focus of assessments because then they would know what was

expected of them before they began a task and could then determine if they had met the expectations when they completed their work.

The incorporation of students' knowledge and background experiences within the context of a supportive educational environment was cited as one of the advantages of authentic assessments. Several informants reported that students gain confidence in their own abilities as learners by taking part in activities such as portfolio assessments. In contrasting the positive experiences of authentic assessment with the pressure of high-stakes tests, one educator described the repercussions of high-stakes tests as devastating. "It is just one more way of showing kids how they have failed".

All of the informants cited time management as the major disadvantage of authentic and portfolio assessments. A number of educators described the difficulties involved in "keeping up with portfolios"; however, most indicated that the effort was "worth it" because of the high quality information obtained about the students' abilities. A bilingual teacher repeatedly praised the value of portfolios in preparing students for high-stakes tests such as the MCAS, but acknowledged that time management was a drawback. Teachers also reported that the storage of materials was difficult and teachers often had to use their own money to purchase the materials required by alternative assessment systems such as portfolios.

SLPs also named time management as a major disadvantage of authentic assessment procedures. Several SLPs reported that although they were cognizant of the value of the information provided by alternative assessments, large caseloads and time constraints imposed by frequent meetings prevented them from completing them

on a regular basis. As a result of the biases of norm-referenced tests, one SLP indicated that she supplements them with curriculum-based and dynamic assessments. Although she favored this multi-faceted approach, the time required in planning and preparing for the authentic assessments was described as being extensive. Several SLPs cited observation as a key component of the assessment process; however, one informant revealed that she sometimes found the organization and integration of the data from the multiple contexts to be challenging. She added that the reports generated by these observations consisted primarily of behavioral descriptors and therefore required a great deal of time to write.

Reflections on the Differentiation of Language Disorders from Language Differences

In this section, the researcher's reflections about the effectiveness of authentic assessment for differentiating language disorders from language differences will be discussed in relation to the assessment procedures. In general, the results from this study found that multiple sources of information were required to differentiate language differences from language disorders. In the majority of the cases data from authentic as well as more standardized measures facilitated in this process.

Social and Educational History

The process of distinguishing language differences from language disorders was usually initiated with a review of students' cumulative records. This first step was vital for providing the framework for interpreting students' current level of

vital for providing the framework for interpreting students' current level of performance and other sources of information, as well as for guiding the topics of discussion in the interviews with students. When examining educational records, it was important to question reporters' level of familiarity with students' language and culture, the processes of first and second language learning, and the phenomena of language loss.

In many cases, older students were present as the records were skimmed prior to the initiation of the assessment. The value of their inclusion was discovered inadvertently when, as a result of time constraints, a cumulative record was perused in the presence of a fifth grade student. The spontaneous dialogue that developed between the researcher and the student provided such an interesting perspective of the student's educational history that this process was replicated whenever feasible. Obviously the confidentiality of certain records was closely guarded during these interactions.

Language Proficiency and Level of Acculturation

The perusal of the cumulative record led naturally into interviews with students about their language use in the home, school, and neighborhood settings. When discussing students' past educational experiences, it was necessary to probe for specific information about the language(s) used in students' previous educational placements. Other areas that required investigation included discussions about former teachers' proficiency level in Spanish and English. For example, one student's records indicated that he had been enrolled in a bilingual program; however, he

reported that his teacher spoke "a little bit of Spanish". The ratio of time devoted to instruction in each language in individual classrooms was another topic that warranted exploration during student interviews. A number of students were enrolled in bilingual programs; however, English was reported and/or observed to be the primary language of instruction.

The investigation of the language usage patterns in students' previous educational placements was supplemented with the assessment of students' level of acculturation. The determination of students' level of acculturation was found to be a vital factor in interpreting results from formalized assessment and in guiding the language(s) of assessment (Gonzales & Kayser, 1997).

The adaptation of the SASH-Y in which three choices were offered (A= Spanish, C= Both, E= English) identified baseline levels of acculturation; however, it did not discriminate the more subtle levels of acculturation required for distinguishing language differences from language disorders. This information was crucial when attempting to evaluate the communicative skills of younger students who had been exposed to both languages. As a result of their limited metalinguistic abilities, these students were particularly difficult to interview in a reliable manner. The pictorial adaptation of the SASH-Y previously described was especially useful with these younger students who were usually capable of providing reliable information about their patterns of language in relation to a familiar individual. The comparison with this familiar individual served as a point of departure for the follow-up questions. For example, if a student stated that she spoke both languages "like her teacher", additional qualifying information was sought (Do you speak Spanish more often or

English more often?). When used in combination with family member and/or teacher report, the pictorial adaptation of the SASH-Y provided a means for obtaining a more precise degree of information about students' level of acculturation.

Results from the SASH-Y were frequently supplemented with interviews that explored issues of students' linguistic abilities and language preferences. Because of the emphasis on learning English in school, many of the students reported that they spoke better English than Spanish. Careful questioning revealed that many of these students were in the initial stages of learning English. In cases in which a record review indicated that the student was probably more proficient in Spanish, yet insisted that he/she usually spoke English, the introduction of a challenging task in English oftentimes resulted in more accurate reporting behaviors. The observation of more subtle behaviors substantiated initial impressions. For example, one student who insisted that he always spoke English was observed to speak consistently in Spanish when clarifying responses not fully understood by the examiner. In another case a student, who attended a bilingual classroom in which English was the primary language of instruction, reported that she spoke English more often than Spanish. As the student's mother had described her daughter as Spanish-dominant, the reliability of her responses was questionable. When told that the tasks would be completed in English, she quickly modified her response.

The determination of students' language usage patterns, proficiency levels, preferences, and degree of acculturation were essential information for the appropriate interpretation of assessment results. This information helped to differentiate those

students who evidenced difficulty on linguistic tasks due to factors such as language loss or second language learning from those who were disordered.

Information from the background history supplemented results from the SASH-Y. Given the relationship between generation level and acculturation (Gonzales & Kayser, 1997), the birthplace of the student and parents was determined through the examination of students' birth certificates that were usually available in cumulative records. The grandparents' place of birth was more difficult to ascertain. Older students were questioned directly, while younger students were asked where their grandparents lived or which language they spoke most often. The socioeconomic status of families was also verified through an examination of students' cumulative records. Because a number of the parents were not employed in positions for which they had been trained, attempts were also made to identify their level of education.

When combined with a record review and interviews with the teacher, student, and parents, the completion of an acculturation scale facilitated in the substantiation of language differences from language disorders. First, it provided baseline information for interpreting the influence of students' linguistic experience and thus determined the primary language of the assessment as well as the language that served as a facilitator.

Second, the assessment of acculturation level provided a frame of reference for interpreting the influence of linguistic and cultural factors upon students' level of academic performance. It also helped to document the influence of language loss and the process of second language learning on norm-referenced tests, especially those administered by other members of the assessment team.

Interviews

Nelson's (1994; 1998) concept of using interviews to identify the "zones of significance" or the specific aspects of the curriculum that are problematic for the student facilitated the differentiation of language disorders from language differences. The triangulation of information from the student, parents, and teacher(s) helped determine the presence of a disorder. Interviews with students about their perspectives of their ability to access the classroom curriculum were also used. In the overwhelming majority of cases, students accurately analyzed their strengths and areas of need. Many older students evidenced their metacognitive and metalinguistic abilities through the identification of the specific cueing techniques and classroom accommodations that facilitated their learning.

Results from the completion of the Listening/Understanding/Remembering Interview (see Appendix H) were especially interesting. Not only did these interviews provide a perspective for the researcher of students' ability to meet the demands of the educational context, but they were also a source of information about strategies that facilitated students' abilities to access the classroom curriculum. Interviews were conducted with eight students who were enrolled in the second grade or above and were reported to have auditory comprehension difficulties. In the majority of cases, older students described a wider array of strategies that facilitated their classroom performance and a higher level of metacognitive awareness of when and how these strategies worked best. Younger students and those with more substantial difficulties tended to rely on assistance-based strategies such as asking the teacher or a peer for

help. In one case an older student who had received supported education services for a number of years in another state, evidenced a high level of awareness of the strategies that were most helpful for her and those that had been attempted and had not worked. This information assisted in the implementation of specific classroom accommodations that facilitated her learning.

Teacher interviews identified students' current performance level as well as their strengths and areas of need in comparison to other students, particularly those with similar educational, cultural, and linguistic experiences. Whenever feasible, interviews with several teachers, past or present, were completed to corroborate the findings. In most cases, the information elicited from the teacher interview was substantiated by students' responses. Although the zones of significance identified by the teacher and student directed the aspects of the curriculum to be addressed in the evaluation, the data obtained from this study supported the value of retaining a comprehensive view of the communication process. For example, a teacher in the monolingual program referred a student for a lateral lisp; however, the completion of a miscue analysis revealed severe word retrieval difficulties. The inclusion of an integrative assessment procedure revealed a communicative disorder that affected the student's educational performance, specifically in the area of oral and written tasks.

This study found the completion of parent interviews to be a vital component in the differentiation of language disorders from language differences. Parents were key informants who provided data about their child's developmental history, the family history of specialized learning or communication needs, and descriptions of concerns about the child's communicative skills. This study was in accordance with

Restrepo (1998) who concluded that a detailed parental interview was a valuable source in the identification of students with communicative disorders. These finding were supported by Gutierrez-Clellen and her colleagues (2000) who concluded that one way for SLPs to differentiate language disorders from language differences was to ask parents to compare the communicative abilities of the child with those of siblings or other peers in the family.

Formal Tests

Despite the major criticisms of norm-referenced tests, this study supported Nelson (1998), who developed the idea of curriculum-based language assessments, when she commented that formal assessments were useful for diagnosing language disorders and determining eligibility for service. When used in conjunction with observation, interviews, dynamic assessments, and curriculum-based language assessments, the judicious inclusion of norm-referenced assessment procedures was found to facilitate in the identification of a communicative disorder, especially for the more covert aspects of language such as comprehension. Multiple sources were found to be essential for the differentiation of language disorders from language differences as each data source balanced the information derived from the other. The assessment of dynamic and multifaceted phenomena such as language required a comprehensive and multidimensional approach to its evaluation.

The pairing of standardized tests and authentic measures was found to facilitate the differentiation of language disorders from language differences. For example, a norm-referenced test was administered to a third grade student revealing his

performance level to be within age-expected levels. His teachers had questioned his ability to process auditory information. While completing a miscue analysis, he was observed to read for meaning construction for the first 10 minutes of the task, at which point he showed a significantly decreased level of attention to the task. In his retelling, he evidenced difficulty in the recall of specific events that had occurred in the story. These behaviors corroborated information obtained during the interviews with his teachers and his parents. A communicative assessment confined to norm-referenced assessments would not have substantiated these behaviors that adversely impacted the student's educational performance.

In another case, a Spanish-speaking student entered the bilingual program with an IEP based upon results from English assessments. The student was assessed on an ongoing basis within the therapy setting. Based upon documentation of his responses in the group setting (e.g., answers to questions in the interactive story reading and the individual conference), he was judged to present with significantly decreased auditory comprehension skills. Following the completion of formal testing, the student was found to have auditory comprehension skills that were within age-expected levels; however, he required the provision of additional wait time. Thus, the inclusion of objective data into the assessment process altered the researcher's clinical impressions of this student's learning needs.

Another Spanish-speaking student arrived in the local school system with an out-of-state IEP. When provided with an at least moderate level of cueing within the dynamics of the group setting, she provided a high level of "I don't know" responses. When tested on an individual basis, she was found to present with no more than a mild

communication disorder. When questioned about her limited level of participation in the therapy setting, she reported that she "didn't want to answer the questions". In both of these situations, formalized testing procedures qualified the findings obtained from more naturalistic communicative settings.

IDEA-97 requires that after identifying the presence of a disabling condition, evaluators must determine whether the condition adversely effects students' educational performance. The findings from this study strong agree with Merrell and Plante (1997) who concluded that norm-referenced tests were not useful for determining the impact of a disability on a student's ability to access the curriculum. The linkage of language assessment to the classroom curriculum was found to be mandatory for this determination. As voiced by many of the interviewed educators, a balanced non-biased approach to the assessment process must include multiple sources of data from both formalized and authentic assessment procedures. The study found that norm-referenced tests must be supplemented with other sources of information in order to obtain a holistic view of the student as a communicator in the school setting, as required by IDEA-97.

Early Literacy Assessments

On the Concepts About Print subtest of the Observation Survey (Clay, 1993), several students with auditory comprehension problems demonstrated difficulty in precisely pointing to the printed words read by the researcher. A consultation with a Reading Recovery Teacher/SLP revealed that these behaviors were often evidenced by students with auditory comprehension difficulties, specifically in the perception of the

individual words contained in spoken sentences (M. Carroll, personal communication, March 1, 2000). Following this consultation, the ability of the students to precisely match voice to print was reassessed. With the exception of one student, all demonstrated a significantly improved performance level on this task. Their improved ability to match voice to print was considered to be the result of their exposure to these activities. These findings demonstrated the importance of assessing phonological awareness and other literacy-based skills in a dynamic manner, especially in cases where students may have had limited exposure to print prior to their enrollment in school. However, given the potential impact on academic performance, these observations of the relationship between auditory comprehension difficulties and the ability to match voice to print should be considered when determining students' eligibility for services.

Running Records and Miscue Analyses

The investigator experimented in the analysis of students' language use with both running records and miscue analyses. An advantage of running records was that they could be used with any student at any time using any type of written material including classroom and expository texts. Because the preparation of the transcript was not required beforehand, running records were extremely flexible and could be used in cases of unplanned assessments. However, because the text was not transcribed for running records, the quality of the language used by students could not be effectively analyzed within the context of the material that had been read. This was

found to be especially problematic when reviewing the material at a later date or when presenting the assessment results to parents and/or teachers. Although a summary of the results of the miscue analysis was filed in students' communicative/literacy portfolios, the lack of context for the analysis had the potential to mask important patterns of behaviors easily discerned through an examination of the actual transcript.

In order to support the development of reading strategies, students read texts at the 90% accuracy level (Clay, 1993). Strict adherence to this accuracy level was found to be problematic for students in the process of English language learning. Many of the students demonstrated miscues such the omission of inflectional endings (e.g., walk/walks), inappropriate addition of inflectional endings (e.g., runned/ran), omission or incorrect use of articles (e.g., the), or inappropriate use of prepositions (e.g., walk for the street/walk in the street). Although a number of these students read texts at accuracy levels well below the 90% level, an adequate level of understanding was evidenced in their retellings. Flexibility in the assessment of language and reading skills was needed when working with students in the process of learning English. The use of surface level criterion such as reading accuracy at designated levels had the potential to seriously underestimate the reading abilities of these second language learners.

Spanish is a language with high orthographic transparency; that is, it demonstrates a close sound-letter correspondence (Valle-Arroyo, 1996). As a result of its transparency, some students read Spanish with 90% accuracy or better, but exhibited very limited understanding of the material. At times, the analysis of surface skills such as reading accuracy levels overestimated students' reading abilities in

Spanish. Running records were not found to be a comprehensive means of assessing Spanish-speaking students' use of language skills while reading aloud.

Clay's recommendation that students read material at a 90% accuracy level yielded a limited number of miscues. It is through the analysis of miscues that conclusions were drawn about students' use of language for the construction of meaning from a text (Weaver, 1994). As a result, the limited number of miscues obtained in running records decreased the comprehensiveness of the assessment.

In running records, all uncorrected substitutions, omissions, and insertions are counted as errors; however, Goodman, Watson, and Burke (1987) considered all departures from the text as miscues, including those that have been corrected. The information gained through the analysis of self-corrections provided insight into students' use of language and their ability to monitor comprehension (Weaver, 1994). Because self-corrections were not included in the analysis of running records, this valuable source of information about readers' language strategies was lost.

Because the fundamental purpose of running records is to code, score, and analyze precise reading behaviors, their applicability in the assessment of linguistic skills while reading aloud was limited. Miscue analyses were found to be a more suitable for the purposes of this study.

Researchers in the field of miscue analysis have found that readers use their knowledge of what is phonologically, syntactically, and semantically possible in a given language to make predictions and anticipate meaning in written texts (Goodman, et al., 1987; Hudelson-Lopez, 1977). Miscue analyses combined with story retellings were found to be an ideal means for assessing the dynamic relationship between the

language user and the written text. Goodman and her colleagues (1987) argued that the degree to which miscues produce structures that sound like language (syntactic acceptability) and make sense within the context of the story (semantic acceptability) reveal students' level of reading proficiency. These authors noted that skilled readers tend to self-correct miscues that disrupt meaning. Observations completed as part of this study found these conclusions to be applicable to the reader as a language user.

When texts were presented that were within students' "zone of proximal development" (Feuerstein, 1979), they clearly evidenced their syntactic and semantic knowledge. Analysis of syntactic miscues revealed language users' ability to understand and produce the structural aspects of language at the sentence and text levels. For example, a second grader demonstrated his knowledge about the way in which object pronouns are used in Spanish when he read "¿En que puedo servirte?" the more familiar form, for "¿En que puedo servirle?" (How can I help you?), the more formalized version that appeared in the text. This student continued to exhibit his knowledge of language during his oral reading when he subsequently anticipated the singular form of a verb, but immediately self-corrected his miscue when he realized that the plural pronoun that followed required a plural verb. In cases where students' substitutions resulted in non-words, their miscues retained the grammatical features of the original word such as "inmedimente" for "inmediatemente" (immediately). According to Weaver (1994) miscues such as these demonstrate students' knowledge of the grammatical constraints of the language.

Rivera-Viera (1978) found that many of her subjects demonstrated semantic miscues when they substituted a more familiar word for the one that appeared in the

text. Similar behaviors were also exhibited by many of the students in the present study. Several students who read the same text substituted "canasta", a word used more frequently by students in this area than the "cesta" (basket) that appeared in the text.

According to Goodman, Watson, and Burke (1987), the use of appropriate stress is related to knowledge of syntax and provides information about the reader's ability to process language. Rivera-Viera (1978) found that the intonation variable was more important for Spanish-speaking students reading in Spanish than in English. In Spanish, within word patterns of stress are important as changes in stress, usually conveyed through accents, can change the meaning of words. In some cases, the grammatical function of a word changes as the result of a change in stress (Rivera-Viera, 1978). For example, "un camino" is a noun that refers to "a road", while "caminó" is a verb that means "he walked". A change is stress conveyed through the use of an accent can change the verb tense of a word. For example, "camino" is in the present tense while "caminó" is in the past tense. The use of stress marked through the use of accents changes the meaning of a word such as "ésta" meaning "this" and "está" meaning "he is". Because the rules regarding the use of stress are highly regularized in Spanish and exceptions to the rule are indicated through the use of accents, the inappropriate use of stress in words can create a non-word such as musíca for música (music).

In her study, Rivera-Viera found that for a group of non-skilled readers, the improper use of stress in Spanish created difficulty in understanding the meaning of the material. Similar results were found in this study. Several students with moderate

and moderate-to-severe auditory comprehension deficits, were very fluent decoders. These students could decode texts that were several grade levels above their instructional level; however, an analysis of their retellings revealed a very limited understanding of the material. It was observed that many of these students, who frequently read the material at a very rapid rate, used inappropriate within-word stress patterns when reading in Spanish. In many cases, they did not pause or attempt to correct their intonational miscues that resulted in non-words. Students who evidenced adequate or mildly decreased listening comprehension difficulties, paused or attempted to correct the non-words created by their intonational miscues. Although they were not always successful in the correction of their miscues, story retellings and/or question probes revealed a fair-to-good understanding of the text. Students with a significant level of auditory comprehension deficits and strong decoding skills self-corrected or attempted to self-correct these intonational miscues far less frequently then the students with mildly decreased comprehension skills, suggesting that the former were not reading for meaning.

Nelson (1998) noted that miscues demonstrated during oral reading identified those areas of oral communication that were not fully developed. Findings from this study strongly supported this premise. Restrepo (1998) found that article errors for gender agreement were one of the most common types of errors in Spanish-speaking children with language disorders. In this study, students who evidenced difficulties with article-noun agreement during oral language tasks also demonstrated these errors during the oral reading as well as the retelling components of the miscue analysis. In

one case, 91% of a student's miscues involved article-noun agreement such as "Yo quiero un bicicleta" for "Yo quiero una bicicleta".

Restrepo (1998) suggested that because Spanish is a more inflected language than English, the analysis of noun phrase errors contributed to the differentiation of language differences from language disorders. Students who demonstrated difficulty with noun-verb agreement in their oral language production also evidenced these miscues while reading aloud. For example, one student substituted "Ellos compró" for "Ellos compraron" (they bought). Miscues involving noun-verb agreement occurred far less frequently than those involving article-noun agreement.

A qualitative analysis revealed a high correlation between the linguistic behaviors represented by students' oral reading miscues and those contained in their oral language production. In addition, several students with histories of expressive language deficits, but whose oral language no longer contained these errors, demonstrated a recurrence of difficulty in the use of these linguistic structures during oral reading tasks that were slightly above their instructional reading level. As a result of the integrative requirements of the oral reading task, the miscue analysis was found to be an excellent source of information about the integrity of students' linguistic systems.

Story Retellings

Findings from this study supported the contention that once readers attain fluent decoding abilities, the processes between reading and listening comprehension become closely related (Snow et al., 1998, 1999). Without exception, all of the

students with auditory comprehension deficits evidenced commensurate reading comprehension difficulties. Because Spanish is a language with a high level of orthographic transparency (Valle-Arroyo, 1996), some students who had mastered the system of sound-letter correspondence were able to read texts with fluency, but demonstrated a limited level of comprehension. These students' retellings evidenced misinterpretations and the absence of important details. Although one second grader was able to decode at the second grade level, his comprehension of the material was at the kindergarten or early first grade level (Beaver, 1997). To differentiate students who could not demonstrate their understanding of the material because of expressive language difficulties, from those who did not understand it, students were asked to summarize the stories in picture form. Although the process of drawing appeared to facilitate some students' recollection of the story's events, other students continued to exhibit a limited understanding of the story.

Gillam and Carlile (1997) analyzed the oral reading and story retelling abilities of students with specific language impairments. These authors found a higher percentage of story retellings rated as "confusing" in students with language disorders. These story retellings included additional events that were not contained in the original stories and/or events that were incorrectly sequenced, and therefore misled the listener. Results from this present study supported these authors' findings. The majority of students with language disorders demonstrated story retellings that were highly disorganized. In several cases, students evidenced an adequate level of understanding of the story; however, their retellings were so poorly organized as to be nearly incomprehensible to the researcher. Students often began their retelling with

events from the final portion of the story and when probed to tell "what else happened", returned to the beginning of the story. In many cases the chronological order of events that took place in the story were incorrectly sequenced.

In their study, Gillam and Carlile's (1997) also noted that the story retellings of the students with language disorders contained information that was not present in the original story. This was the second most common characteristic of the story retellings of students with language disorders evaluated for this study. In most cases students provided information that was marginally related to the original story suggesting that they had understood isolated components of the story, but had misinterpreted the majority of the events. For example, in a retell of a story about a chicken who wanted a bicycle for her birthday, one student indicated that it was the woman's birthday and then related a story about that aspect of the story. Other students were observed to rely primarily on the information contained in the pictures.

In several cases the comprehensibility of the students' retellings was decreased due to the lack of referential cohesion. Referential cohesion, or the way in which characters, places, or props are introduced, facilitates the listener's understanding of the events that occurred in the narrative (Gutierrez-Clellen & Heinrichs-Ramos, 1993). In stories with several characters or props, students with moderate language disorders provided an insufficient level of information about which object, person, or place they were referring to in their retellings. For example, after reading a story about two friends who participated in a variety of separate activities after an argument, a second grader described the characters' actions, but did not clearly convey which character had engaged in which activity. Students with milder linguistic impairments or speech

disorders (articulation or fluency) relied on strategies such as the use of elliptical reference to explicitly state their ideas. For example, one student first introduced the main character by saying, "El niño no encuentra su gorra" (The boy couldn't find his cap). Because he had already identified the main character, the student's listing of the character's actions was understood by the listener.

This study found the use of story retellings to be an excellent way to assess students' integrative language skills for the completion of discourse level tasks. In addition, results from the task corroborated and supported information derived from other sources such as narratives.

Narratives

The rubrics located in Appendix H were found to be highly useful for analyzing the narratives generated by students while viewing wordless books as well as in the story retellings that followed the oral reading tasks. In general, the narratives generated while viewing the visual stimuli were more organized and cohesive than those told following the oral reading. Because the later narratives were a component of an integrated linguistic activity, they were found to be more representative of students' language behaviors as evidenced within the classroom. Students' verbal output from the conversational language samples, narratives, and story retellings was analyzed for patterns of strengths and weaknesses and their impact on students' academic success.

Language Samples

In a recent article, Gutierrez-Clellen and others (2000) addressed the complex issues associated with the collection and analysis of a non-biased language sample with monolingual and bilingual Spanish-speaking children. The processes for differentiating language differences from language disorders used in this study were in compliance with the recommendations offered by these authors. The first suggestion outlined in their paper was that evaluators complete structured interviews with parents, caregivers, or teachers to obtain information about students' language proficiency and patterns of exposure to one or both languages. This topic was previously addressed in this study.

The authors next suggested that evaluators describe students' use of language and tentatively compare it to children of similar sociolinguistic backgrounds. Restrepo's (1998) recommendation that SLPs determine the presence of a significant number of grammatical errors per TU in one language for monolingual children and both languages in bilingual children was found to be valuable in identifying students with language disorders. However, as in all aspects of communication for students from culturally and linguistically diverse backgrounds, these data needed to be interpreted in context. This study found that comparing students' level of acculturation with language sampling facilitated the differentiation of languages disorder from language differences. This pairing expedited the determination of the applicability of the available developmental norms on students, especially those who had been exposed to both languages in the home or through intensive exposure via a daycare or preschool program. Information about students' level acculturation and

family history of special needs in combination with the analysis of language samples helped to distinguish atypical language production errors that were a result of language shift from those that were the result of a language disorder. For students who had been exposed to both languages in varying patterns (e.g., one language in the birth home, another in the foster home, and both languages in the adoptive home), the analysis of level of acculturation helped to select the measure of sentence length (MLU or MLR).

Language and Literacy Rating Scales and Checklists

The CRLAI (Worcester Public Schools, 1999) was initially used to identify the presence or absence of the described behavior (e.g., Does the student use active listening skills?); however, this format resulted in a paucity of information about students' linguistic abilities. When the wording of the statements was altered to an open-ended format (e.g., How does the child use active listening skills?), a more detailed response was obtained about students' ability to use language within the framework of the curriculum. In completing this interview, teachers were asked to compare students' abilities to those of their peers with similar cultural and linguistic experiences.

The Classroom and Communication Learning Checklist (CCLC) (Wiig & Secord, 1994) was used to assess five students with complex social as well as educational histories. Because many of these students manifested multiple deficit areas, the level of concern expressed by the teacher about specific behaviors assisted in prioritizing those areas to be addressed in the assessment. In one case, several

teachers completed this form independently to triangulate the pattern of the student's behaviors across educational contexts. This information was corroborated by more formalized assessments, curriculum-based language assessments, and observation of the student in the bilingual and ESL classrooms.

In cases where a comprehensive evaluation could not be completed prior to the Team Meeting, the CCLC served as a structured interviewing device that covered a wide range of behaviors relevant to success in the classroom. Based upon the suggestion of one of the interviewees, the level of concern voiced by the teacher was converted into a narrative description of the reported behaviors. Specifically, areas of strength were reported for behaviors that never occurred, minimal difficulty was reported for behaviors that rarely occurred, moderate difficulty was reported for behaviors that frequently occurred, and significant difficulty was reported for behaviors that always occurred. For example, according to the classroom teacher, the student demonstrates significant difficulty in spelling words correctly and writing things in a different way; however, he is very organized, always comes to class prepared, and follows directions with minimal difficulty.

The abbreviated version of the Bilingual Classroom Communication Profile (BCCP) (Roseberry-McKibbin, 1995) was utilized in three cases in which a bilingual child was referred for cognitive and communicative testing by a teacher from a monolingual program. Although the BCCP had been used successfully in the past by the researcher, its utility in these cases was marginal. Although services were not recommended in all three cases, the teachers reported a high number of behaviors associated with speech-language deficits, such as the student has difficulty paying

attention even when material is understandable and presented using a variety of modalities (Roseberry-McKibbin, 1995). These referrals were initiated by teachers who worked in schools attended by students from diverse cultural and linguistic backgrounds; however, a bilingual program was not housed in these buildings. These students also presented with complex social and educational experiences that complicated the identification of a communication disorder. The teachers' lack of familiarity with the process of second language learning may have impacted the utility of the BCCP in these three cases.

The Observation of Reading Behaviors, an adaptation of the work of Fountas and Pinnell (1996), was used as part of the oral reading assessment. It directed the investigator's attention to the key behaviors associated with the development of early literacy skills. The increased knowledge about early reading behaviors such as the relevance of the ability to remember and use language patterns or to use oral language in relation to the text, provided a framework for assessing students' ability to access the school curriculum. A recent interpretation of the mandates of IDEA-97 by ASHA (2000) supported the need to include the parameters of literacy as part of a school-based SLP's speech and language eligibility criterion.

The Early Identification of Language-Based Reading Behaviors (EILRB) (Catts, 1997) was used dynamically as a component of the kindergarten screening (can the child remember the majority of classmates' names by mid-October?) to triage those students to be evaluated for speech-language or other supported services. It also served as a tool for validating continued speech and language services for those

students who were demonstrating adequate progress in the classroom, but who were at risk for academic difficulties.

A surprising outcome of this research study was the overall limited applicability of checklists for the differentiation of language differences from language disorders. The Observation of Reading Behaviors alerted the researcher to the early reading behaviors that required careful observation, while other checklists (Catts, 1997) were most useful for the validation of services for students already identified as language disordered. Although the curriculum-based tools such as the CRLAI were useful in maintaining the focus of the assessment on the language-based academic skills as mandated by IDEA-97; checklists and rating scales were most productive in the assessment of precise linguistic behaviors when combined with the interviewing of informants.

Rubrics

As part of the initial evaluation process, the Student Writing Rubrics (Goddard School of Science and Technology Faculty, 1996) assisted in the analysis of writing samples from the portfolios of 15 students. One of the major difficulties in the use of the rubric as a tool in the retroactive analysis of writing samples was gathering information about the circumstances surrounding their production. Typically teachers and students were unable to provide detailed information about the initial instructions, the level of cueing required by the student or the time frame in which the product was completed. In a number of cases, drafts that demonstrated the process used to generate the product were unavailable.

An adaptation of the MELA-O (1994) was employed on three occasions with students who evidenced first language loss as a result of intensive and prolonged exposure to English. Although these students came from homes in which primarily Spanish was spoken, they evidenced a low-to-moderate or moderate level of acculturation as measured by the SASH-Y. Two of the students had never been enrolled in a bilingual program, while the third student had not attended bilingual classes for two years. When asked to compare their present level of Spanish-proficiency skills to those demonstrated upon their entrance to the monolingual program, all three reported a loss of their communicative abilities. In two cases, the students' parents corroborated the finding of language loss as well as the lack of prior communicative deficits or family history of special needs. Rather than treat their language differences as language disorders, the adaptation of the MELA-O (1994) focused on the students' linguistic abilities, not their communicative difficulties that were the result of sociocultural factors.

Story Conferences

In cases where the evaluation could be completed over a longer time period, students' linguistic skills as observed during the story conference cycle supplemented the data obtained from standardized tests. Students attended several sessions of speech-language therapy with their classmates. The observation of the student in a group session allowed the evaluator to compare the students' communicative skills within three different contexts: a one-on-one testing session, a structured small group setting, and the classroom environment. It also provided the opportunity to assess the

students' language skills dynamically and to document the cueing strategies that facilitated their performance level while completing an authentic task. The supplementary information obtained from the story conference contributed to the more accurate differentiation of language differences from language disorders.

When used as a component of a miscue analysis, the assessment of listening comprehension provided data about students' ability to recall factual and inferential information as well as components of story structure about a complete text. The results from this observation of students' listening skills balanced information obtained from norm-referenced tests that typically use short paragraphs to assess these skills. Because students were required to retell the story, they were unable to rely on the cues provided by targeted questions when formulating their responses. The results from this assessment approach were found to be promising, especially as they supported the findings of Nelson (1998) who advocated the assessment of students' discourse-level language processing abilities.

Determining the Language of Instruction

Eight follow-up interviews were conducted with bilingual and ESL educators concerning the most effective means for determining the language of instruction for bilingual students with communicative disorders. Analysis of their responses in relation to the research question revealed criticisms of the current methods for transitioning bilingual students with communicative disorders, particularly the lack of qualitative measures of the cognitive/academic language proficiency skills required for success in the monolingual program.

The majority of the informants cited the observation of the student in multiple contexts over time as the most valid means for determining the language of instruction. Several ESL teachers suggested that during these observations, educators should note the language used most often to initiate conversation and the language in which the student requested the highest number of clarification checks. Others suggested that the students be observed while completing a linguistically challenging activity to compare the type and level of support required in each language. One informant cautioned that in observing students' use of language in the classroom, the sociocultural climate of the classroom and its influence on students' use of language should be considered. She reported that sometimes students use lots of English because of the high value placed on it by the teacher, not because they are more proficient in the language or feel comfortable speaking it.

In addition to the assessment of students' linguistic abilities, informants described interviews with current and former classroom and ESL teachers as valuable sources of information in the decision-making process. According to one SLP, teachers with knowledge of both languages were especially helpful because they were able to provide information about students' dual language and literacy skills.

Student interviews were also cited as a valuable means of determining the language usage patterns of the home, school, and neighborhood settings. In addition to the language used, informants stressed the need to probe students' perceptions of their abilities in the areas of understanding, speaking, reading, and writing in each language. A SLP reported that she investigates the language in which students feel most comfortable communicating and the reasons why.

Informants also stressed the value of parent interviews for the exploration of the patterns of language use by parents, siblings, and extended family members. One educator reported that she interviews parents about their daily routines and inquires about the language or languages used during these activities. She also requests parents to describe their impressions of the child's language proficiency in the first, second, or both languages. She said that she asks specific questions about the child's ability to use language when being taught a new task or when providing information about something experienced only by the child. Another educator suggested that parents be interviewed about their child's communicative abilities in both languages as compared to other children with similar cultural and background experiences or with students who are presently enrolled in monolingual programs.

Because language samples provide a comprehensive level of information about students' linguistic abilities, they were cited as a valuable tool for determining the language of instruction. Informants commended the flexibility and adaptability of language samples as well as their ease of use when comparing students' skills across languages and learning contexts. Interviewees described an array of elicitation techniques such as retelling a story that had been read aloud, describing the steps necessary to complete a task, or telling a story while viewing sequenced picture cards. An educator reported that she asks her students to relate personal experiences or answer "what-if" questions. An ESL teacher stressed the value of having students discuss topics that are relevant to their lives and personal experiences. The informants recommended that the language samples be qualitatively assessed in both languages to determine the most appropriate language of instruction.

Informants repeatedly argued that higher level linguistic and literacy skills must be considered when determining the language of instruction. A SLP voiced her frustration with cases in which teachers, who lacked recognition of the impact that a communication disorder exerted on the transition process, referred students who didn't have the cognitive/academic skills in either language needed for academic success. She cautioned that care must be exercised to insure that students with communication disorders have requisite the English language proficiency skills to permit their access to the classroom curriculum. An ESL teacher also emphasized the importance of looking beyond oral conversational skills when transitioning students. She cited the critical need to examine students' English cognitive/academic language proficiency skills including their ability to analyze, synthesize, and apply information.

Informants also emphasized the importance of evaluating students' literacy skills in both languages when determining the language of instruction. A SLP reported that she has found the comparison of students' listening and reading comprehension skills across languages to be an effective means for determining the language of instruction. Other educators recommended both the assessment of students' reading skills using curriculum materials and the observation of their responses in discussions of literature. The majority of the informants indicated their preference for qualitative, rather than quantitative approaches for the measurement of reading comprehension.

Work samples were a frequently cited source of information about students' writing skills. Several educators stated that they examine students' portfolios and compare the quality of the work in both languages. A bilingual teacher reported that

she analyzes the rate and amount of improvement in students' work over time to determine the language that has demonstrated the most consistent gains. Another bilingual teacher recommended documenting the level of scaffolding required by students for the completion of work samples as a guide for determining their written language abilities in both languages.

An ESL teacher summarized many of the informants' comments about the standardized assessments currently used to measure English language proficiency. He referred to traditional assessments as culturally inappropriate, surface oriented, and incapable of measuring the complex array of linguistic skills necessary for success in the monolingual program. He added that given the current expectations from teachers and students, a more comprehensive language tool such as the portfolio is required to measure the high standards mandated by the Massachusetts Curriculum Frameworks and the MCAS.

Reflections on the Determination of the Language of Instruction

The complex process of determining the language of instruction for bilingual students with communication disorders was found to require the analysis of multiple sources of information combined with a comprehensive assessment of dual language systems. Qualitative sources of data such as miscue analyses and curriculum-based assessments in conjunction with record reviews, interviews, and measurements of acculturation yielded a more accurate evaluation of students' linguistic and literacy skills in two languages than traditional testing procedures. In order to determine the recommended language of instruction, analogous tasks were chosen in each language

to allow the comparison of students' linguistic behaviors between the two languages. The following represents a description of the observations and conclusion made by the investigator regarding the findings of the most effective means for determining the language of instruction for bilingual students with communicative disorders.

Social and Educational History

As in distinguishing language differences from language disorders, the record review process provided the framework for the accurate interpretation of students' current academic performance as well as their potential to access the curriculum of the monolingual program. Information that contributed to the decision-making process included the length of time the student had been enrolled in the bilingual program, the predominant language of instruction in the previous educational settings, and the current level of performance in the bilingual and ESL classrooms. The examination of the cumulative records, including annual classroom performance checklists and measures of English proficiency skills such as the MELA-O (1994), provided a reference point for comparing and contrasting students' dual language abilities.

Level of Acculturation

Students' level of acculturation was determined through record reviews and the administration of the SASH-Y or its adaptation. The information obtained from the SASH-Y was supplemented with interviews with students about their perceptions of their language and literacy skills in both languages (e.g., which language do you

speak, read, or write best?). Issues related to students' language preferences and the reasons for these preferences were also explored in order to identify sociocultural variables that may have influenced their responses (e.g., a classroom or school where English language learning was strongly emphasized). The results of both the SASH-Y and the interview were corroborated with teachers and family members such as a parent or an older sibling enrolled in the school. The program enrollment of siblings was also investigated as this has been often been observed to be a predictor of students' placement in the monolingual program.

The analysis of students' level of acculturation through the use of the acculturation scale and record review was found to be a key component in determining the influence of linguistic and cultural factors on all measures of academic performance. When combined with the evaluation of their cognitive and academic language skills, the analysis of students' level of acculturation provided the necessary framework for interpreting assessment results that helped to determine the recommended language of instruction in bilingual students with communicative disorders.

Interviews

In addition to the language proficiency and preference interview, specific questions that required students to compare their skills across languages were integrated into the majority of the interviews. For example, on the Listening/Understanding/Remembering Interview (Appendix I), students identified the language in which they remembered things most easily. On the Language and

Reading Observation Guide (Appendix F), students were asked the language that was easiest for them to read and why. For students already enrolled in speech-language services, these topics were discussed periodically throughout the year. As part of initial assessments, questions were selected based on students' age as well as their demonstrated pattern of strengths and weaknesses. For example, questions about the use of specific memory strategies would not be discussed with a student with word retrieval difficulties.

Teachers were also interviewed about students' patterns of language use in the classroom in the areas of comprehension, speaking, and written language. Information was elicited about students' language preferences in a variety of contexts in which the use of either language would be appropriate such as when playing with bilingual peers at recess. Bilingual teachers were asked which language they used with a student in questions when they really wanted the student to understand something, as well as the language that the student used with bilingual speakers when he/she really wanted them to understand something.

Parents or other family members were also interviewed about the student's pattern of language use, proficiency, and preference in the home and neighborhood settings. In cases where parents could not be contacted, a copy of the Language Preference and Home Language Survey Form (Worcester Public Schools, n.d.) was sent home to obtain updated information about the child's language behaviors as viewed by the parents.

Communicative and Literacy Portfolios

In order to provide a frame of reference for comparison of students' performance across languages, artifacts were collected in both languages on an ongoing basis. This practice supported the recommendations of the informants who indicated that one of the best ways to determine the language of instruction was through the analysis of daily work samples. For students who were recommended for transitioning into the monolingual program, the gathering of artifacts in both languages was completed in a systematic manner previously described that allowed the comparison of their language and literacy skills for equivalent tasks. The following is a description of the items contained in students' portfolios and how their inclusion facilitated in determining the language of instruction.

Running Records, Miscue Analyses, and Retellings

When determining the language of instruction, miscue analyses were the preferred method for assessing students' reading comprehension skills. As previously reported, Clay (1993) maintains that students develop target reading strategies most effectively when they are reading material at the 90% accuracy level. Because students who are in the process of learning English frequently demonstrate errors in the use of inflectional endings (e.g., run/runs), articles (e.g., the), and prepositions (e.g., at), the use of a running record had the potential to underestimate their reading abilities. Due to the close sound-symbol relationship in Spanish, a number of students in this study read texts with a high level of accuracy, but with a limited level of understanding of the material. Because one of the purposes of this study was to

investigate ways in which authentic assessments could be used to determine the language of instruction, running records were not found to be an effective means for comparing linguistic skills and reading comprehension across languages.

When comparing results from miscue analyses completed in both languages, this study found a clear relationship between the percentages of miscues that retained meaning and students' level of comprehension of the text. That is, the majority of the students read material in their more proficient language with fewer miscues that compromised meaning and higher levels of comprehension as measured by their retellings. When reading in their more proficient language, students evidenced behaviors associated with ongoing comprehension checks, such as rereading to confirm or using information from multiple sources in an integrated way while focusing on meaning (Fountas & Pinnell, 1996). While reading a story in English about a mother beaver, who placed her life in danger to protect her young, one student demonstrated her ability to predict outcomes when she gasped and inquired if the bear was going to kill the mother. This student clearly demonstrated a high level of interaction with the English text that was confirmed by her story retelling and she was recommended for transitioning into the monolingual program by her teachers.

Goodman and Burke (1973) found that graphophonemic cues were used more often by less proficient readers than by more proficient readers. In applying these findings to reading behaviors in both languages, the majority of students were observed to rely more on graphophonemic cues rather than semantic or syntactic cues when reading in their less proficient language. In addition, this reliance on graphophonemic cues produced a higher number of non-words that were not meaning

preserving at the sentence and text levels. In general, the analysis of the retellings of texts that contained a high number of non-words, revealed an overall understanding of the "gist" of the story, but limited reference to supporting details or inferential information.

The analysis of the behaviors associated with fluent reading of skilled readers in both languages revealed different levels of performance when reading in their more and less fluent languages (Pinnell, Pikluski, Wixson, Campbell, Gough, & Beatty, 1995). When reading in their more proficient language, skilled readers who demonstrated a satisfactory level of understanding read in larger more meaningful phrases, with occasional slow downs at points of difficulty; reread for problem solving; and attended to punctuation (Pinnell, et al., 1995). When reading in their less fluent language many of these same students presented with behaviors associated with a decreased level of understanding of the text, which was confirmed through their retellings. For example, some of these students were observed to read in a syllable-by-syllable or word-by-word manner, frequently pausing between words, and paying a limited amount of attention to punctuation. Others read in a very rapid fashion, rarely pausing to reread for problem solving and paying little attention to punctuation. This study found that skilled readers demonstrated a superior level of reading fluency when reading aloud in their more proficient language.

For example, one student who was performing within grade-expected levels in the classroom and received services for a speech disorder was observed while reading a story in English. He read the text that was one year below grade level at a very rapid rate, frequently ignoring punctuation, and rarely pausing to reread or self correct at

points of difficulty. Analysis of his oral reading revealed that many of his miscues involved omissions of inflectional endings and non-content words, while others resulted in non-words. In his retelling, he evidenced a generalized understanding of the story, but was unable to provide supporting details. Although he read a story in Spanish that was slightly above grade level and at a slightly rapid rate, his miscues preserved meaning. In his retelling, he evidenced a satisfactory level of understanding of both factual and inferential information from the story. A positive relationship was found between the quality of this student's reading fluency skills and the level of understanding of a text read in his more fluent language.

When students read texts in both languages, a comparison of their retellings revealed a higher level understanding of the material in what was presumed to be their more proficient language for academic purposes. These retellings evidenced a superior ability to recall important events, comprehend factual as well as inferential aspects of the story, and organize and sequence the information in a more comprehensible manner. Their use of vocabulary was found to be more specific and precise and contained a wider range of words or phrases from the story. In addition, students showed an increased ability to respond to the text and relate it to their background experiences.

For example, miscue analyses were completed in Spanish and English with three students who demonstrated very fluent oral language skills in English, but who continued to benefit from instruction in Spanish when learning new information. The completion of miscue analyses in English revealed the ability to use surface features such as inflectional endings in their retellings; however, they demonstrated a minimal

level of understanding of factual information from the story and a very limited level of understanding of inferential information. These students demonstrated a significantly superior level of understanding when reading much more difficult texts in Spanish. The organization of their retellings was qualitatively superior in Spanish and the incorporation of words and phrases from the story was noted as well.

Nelson (1998) suggested that oral reading miscues identified those areas of oral language that were not fully developed. This study found this idea to be applicable to the phenomena of second language learning. Although the majority of the students demonstrated errors typically associated with English language learning such as deletions of inflectional ending, those who were in the earliest levels of English acquisition exhibited the highest number of these miscues. Students, who were considered to be strong candidates for placement in the monolingual program without the supportive services of special education, evidenced a higher level of self-correction of these miscues during both the oral reading and retelling components of the task. These students also exhibited adequate levels of comprehension of the material as measured in their retellings. This study found that the analysis of the quantity and quality of students' miscues in English provided insight into their level of mastery of this language.

Ortiz and Garcia (1990) defined one of the characteristics of language dominance as the one whose rules and structures seem to influence the production of the other language most frequently. These behaviors were observed to occur with some of the students in this study. Several children inverted the position of nouns-adjectives while reading Spanish to reflect the word order rules of English. For

example, one student initially read "sombrero amarillo" as "amarillo sombrero", but immediately self-corrected his miscue. While retelling a story about a birthday party, another child demonstrated the influence of Spanish in her verbal output when she stated that the "little boy had four years". These observations contributed qualitative information about the relative levels of proficiency of each language as ascertained by one language's level of influence on the other.

The Oral Reading Observation Guide (Appendix J), an adaptation of the work of Fountas and Pinnell (1996), was highly valuable in qualifying reading behaviors in students who demonstrated a similar level of reading performance in both languages or with older students who demonstrated skills at the very earliest levels of literacy. The ability of this checklist to compare specific reading behaviors between Spanish and English facilitated the decision for the recommended language of instruction. This was found to be an especially critical decision for students who had been enrolled in supported bilingual services for several years, but continued to demonstrate slow progress in their development of literacy skills. For example, an older student who had been taught primarily in English demonstrated reading skills in both languages at the early first grade level. Analysis of his specific reading behaviors revealed superior word recognition skills in English; however, he demonstrated a more efficient use of syntactic and semantic cues in Spanish. It was recommended that he continue to receive instruction in an educational setting that offered him access to both language systems.

Early Literacy Assessments

The Observation Survey (Clay, 1993) served as a means for assessing students' emergent literacy behaviors in English and Spanish specifically in their ability to implement target strategies with words, letters, and texts. Although appropriate for use with first graders and older students who demonstrated skills at the early stages of literacy development, it was especially useful for students at the end of kindergarten. Because many of the referrals for transitioning were based upon these students' oral language skills, Clay's Observation Survey provided a means for comparing students' early cognitive/academic language skills such as their knowledge about print and ability to read and/or write letters and words in each language. The information about students' beginning reading and writing skills contributed to the decision for the language of instruction.

For example, a teacher had recommended that one student who required additional response time be transitioned to the monolingual program. Because the teacher often repeated her questions in English, she misinterpreted the student's slow response rate as responding to the English stimulus. The administration of the Observation Survey revealed a significantly superior performance on the Spanish portion of this instrument when compared to his English performance and the student was not recommended for transitioning.

Language Samples and Narratives

The collection of language samples for a variety of discourse level tasks provided insight into students' ability to use language for social as well as academic

purposes, critical factors related to academic success. As suggested by Nelson (1998) the analysis of the ability to describe discourse level events for conversational, narrative, and expository purposes was found to be an excellent vehicle for comparing students' basic interpersonal and cognitive/academic language skills in Spanish and English. For example, samples were collected during conversations about subjects of interest, while viewing a wordless book or a CD-ROM, and giving directions on how to play a game or complete a specific task.

For students who had been recommended for transitioning, language samples were collected in both languages and the complexity of the output was analyzed through the use of terminal units (Hunt, 1965; Restrepo, 1998), as previously described. The discourse level output was also qualitatively analyzed using the rubric that can be found in Appendix H. A language sample was collected or taped as students interacted in English in the classroom or ESL settings. The focus of the analysis of these language samples was on the students' ability to use language for academic purposes and contributed yet another piece of information to the determination of the language of instruction.

Audio Samples

As a result of scheduling conflicts, several students who had been recommended for transitioning could not be observed in their ESL or English reading classes. Teachers taped these students and then reviewed the tapes with the researcher while providing background information about the taping. This alternative to onlooker or participant observations provided valuable information about students'

ability to use English for authentic purposes within the naturalistic context of the classroom routine.

Rubrics

The Student Writing Rubrics (Goddard School of Science and Technology Faculty, 1996) provided a listing of characteristics for writing behaviors at each grade level. Because it was developed by monolingual and bilingual program teachers, these behaviors were applicable to writing in Spanish and English. For example, descriptors for the kindergarten level included: writes name and favorite words; is beginning to use spaces between words; and is starting to read his/her own writing. This rubric served as a means for the qualitative analysis of students' writing skills in both languages.

As previously stated, the Holistic Scoring Rubric for Writing Assessment with English Language Learning Students (O'Malley, 1996) was developed by ESL teachers specifically for use with students who are in the process of learning to write in English. The rubric stressed the organizational components of writing, an area of need for many students with language disorders. However, it did not penalize them for errors in mechanics, an area of difficulty for many English language learners. When used in conjunction with the Student Writing Rubrics, it offered a qualitative means for judging students' writing skills in comparison to students enrolled in the monolingual program as well as those who were English language learners.

The qualitative analysis of students' English proficiency skills as recorded by ESL teachers on the MELA-O (1994) provided insight into students' ability to

comprehend and produce language for interpersonal and classroom discussion purposes. When combined with measures of students' cognitive and academic language skills, it contributed to the determination of the language of assessment.

Story Conferences

Story conferences provided a means for comparing students' understanding of text level material and story structure in both languages. The students' ability to participate in the interactive group discussion and the story conference provided a means for assessing their ability to integrate form and content in each language in response to questions about text level material that had been read to them. In addition, the application of dynamic assessment principles in the form of cueing strategies allowed the comparison of the level of scaffolding required for securing a response to the target questions in each language. Analysis of students' responses to the story conference task found that in their more proficient language, students recalled more factual as well as inferential information about the text, required a shorter response time to answer the questions, and needed a lower level of cueing.

For students who had been recommended for transitioning, their ability to retell the events of a story that had been read aloud was assessed in both languages using texts of similar length and level of difficulty. The students' level of understanding of the factual and inferential aspects of the story was analyzed using the criterion established for the retelling component of the miscue analysis (see page 188). The comparison of students' listening comprehension abilities in both languages using texts of similar levels of difficulty helped to identify the language best understood by

the students. Their performance on these tasks was used to help determine the recommended language of instruction.

Therapy Notes

Therapy notes were used to record observations of students' behaviors in both languages in the therapy and classroom settings. Summaries of parent and teacher conferences and any other pertinent information related to the recommended language of instruction were documented as well.

The determination of the language of instruction for bilingual students with communicative disorders required the analysis of multiple measures of qualitative information. Authentic assessment techniques, especially those that were related to the classroom curriculum, helped to determine if students possessed the requisite cognitive and academic language skills needed for success in the monolingual classroom.

Documentation of Progress

The majority of the interviewed educators characterized authentic assessments as a tangible means for documenting students' progress. The inability of standardized assessments to accurately reflect students' progress were criticisms voiced by many of the informants, who acknowledged their strong support of classroom-based measures such as portfolios and authentic assessments, especially for use with students with special needs or from culturally and linguistically diverse backgrounds. Comments

such as "high-stakes tests do not and cannot accurately measure progress, especially for students with specialized learning needs" were common. A preschool teacher described portfolios as "developmentally based and geared to the individual learning needs of the child, and therefore capable of capturing the gradual progress evidenced by many children". Several teachers contrasted the ability of authentic assessments to show both the process and products of students' learning with standardized assessments that focus exclusively on the products of learning. An ESL teacher voiced her frustration with high-stakes tests when she stated that she feels defeated because she "knows how far her students have come and you can't prove it with standardized tests that just don't show their growth".

According to a bilingual teacher, portfolios and other authentic assessments precisely measure students' achievements on an ongoing basis. She added that authentic procedures allow her to evaluate her students in an individualized way so that she knows what they have achieved and what she needs to do to make sure that they continue to achieve. In praising the ability of students' reflections about their work to accurately document their ongoing progress, she characterized the use of authentic assessments and portfolios as creating "a win-win situation for everyone-students, teachers, and parents!"

As reported by a special educator and seconded by a student who was present during the interview, students "LOVE to look at their work and see how far they have come!" Students' participation in their own learning that led to an increased level of self-esteem and self-confidence was a frequently cited benefit of portfolio assessments. In describing the significantly improved level of achievement evidenced

by one of his students, a bilingual teacher proclaimed that the student's "self-esteem has gone through the roof. He has done a 360 degree turn around!"

Portfolios and authentic assessments were reported to be a valuable tool when included in parental conferences. Many of the informants referred to the merit of tangibly demonstrating the progress evidenced by the child. A bilingual teacher indicated that "parents love to look at their child's portfolio and sometimes are really surprised at what their child has accomplished". A preschool teacher indicated that portfolios clearly demonstrate to parents "if a child is moving along or really struggling". She added that she thought that it was important for parents to have concrete evidence of the child's current level of performance, especially when a teacher wanted to refer the child for specialized services. Another teacher reported that she found it beneficial to provide evidence of the child's growth, especially in cases where a child was having difficulties in other areas. "Parents are more willing to hear about areas of need when you have balanced them off with good news".

Reflections on the Documentation of Progress

The findings from this study echo the sentiments of Linda Eidman (cited in Weaver, 1994) who likened the documentation of progress through the use of standardized tests to the portrayal of a trip around the world with a single photograph. IDEA-97 mandates the development of IEPs with measurable annual goals that support students' progress in the general education curriculum. Authentic assessments were found to be an ideal means for documenting students' progress in the attainment of speech and language therapy goals in the context of the classroom curriculum.

Communicative and Literacy Portfolios

Communicative and literacy portfolios were found to be an effective means for documenting students' progress toward the specific speech and language targets of therapy as well as the impact of the communication disorder on students' educational performance. Kratcoski (1998) maintained that one of the greatest strengths of portfolios was that they demonstrated learners' achievements through the provision of dated items that had been collected at regular intervals over time. This study found that it was the completion of specific tasks at regular intervals that most clearly documented students' progress toward the acquisition of targeted therapeutic goals. Based upon suggestions made by Kratcoski (1998), the ongoing analysis of the ways in which these artifacts reflected communicative growth and achievement was also required. For example, a brief description of a student's ability to answer questions during a group discussion of holiday plans was noted on a "post-it" that was placed on his/her self-portrait. In other cases, a student's ability to address the topic of a target question or use article-noun agreement in written output was analyzed. A summary of these behaviors was attached to the artifact and documented students' ongoing progress in therapy.

As previously discussed, students periodically reviewed their portfolios and reflected on their progress. Students in kindergarten and first grade typically responded to their work in an affective manner. The majority of these students responded to all of their artifacts in a highly positive manner, rating them as deserving a "happy face". Some of the students commented on improvements in the aesthetic

features of the worked samples such as "prettier writing" or "nicer coloring". Slightly older students were able to rate their work samples in a holistic manner. For example, several commented that their artifacts showed that they were able to write longer stories or knew how to write in cursive. When asked to identify what they had learned over the course of the year, one student reported that he was able to read books without pictures and write stories. However, his focus on the aesthetics of his work products was evidenced during preparations for the portfolio conference. He had written a detailed piece on a trip to New York, but did not have time to illustrate his story. He quickly discarded this story in favor of one with a "pretty picture".

Students' active participation in their own learning through portfolio conferences and the evaluation of their work samples was found to be a very positive outcome of this study. Students' reflections on the outcomes of their own learning clearly demonstrated their metacognitive growth and contributed to the holistic evaluation not only of their communicative abilities, but of their linguistic gains within the context of the classroom curriculum.

Story Conferences

Story conferences were found to be an excellent tool for documenting students' progress toward their therapy goals, such as in the ability to understand and recall text level information; to address the content of questions; to discuss the components of the story's structure; or to use specific grammatical forms. Because the stories were often selected based on students' communicative needs, therapy activities were closely tied to therapy goals. The individualized format of the story conference allowed the

researcher to probe the level of support that elicited the production of the target form. Because story conferences were usually presented several times during the school, students' progress toward their therapy goals in the context of the regular education curriculum was closely monitored and documented over time.

Early Literacy Assessments

Clay's Observation Survey (1993) provided ongoing documentation of students' progress in their acquisition of early literacy skills, not only for the researcher, but also for their teachers, parents, and the students themselves. For example, on the concepts about print portion of the Observation Survey, some students demonstrated their increased knowledge of metatextual concepts such as the comprehension of left, right, first, last, letter, and word. Other students exhibited the improved ability to match print to the words contained in sentences read aloud by the examiner. As mandated by IDEA-97, evidence of these linguistic achievements was directly linked to the classroom curriculum.

Miscue Analyses and Retellings

Because miscue analyses required the ability to use language for integrative purposes, they were found to be an excellent tool for measuring and documenting students' progress in therapy. Miscue analyses evaluated an array of language comprehension and production parameters that were directly related to students' progress in the regular education curriculum as mandated by IDEA-97. In addition,

miscue analyses measured students' progress in the mastery of those aspects of language form, content, and use typically associated with speech and language services. The inclusion of miscue analyses in the ongoing assessment process increased the level of collaboration with teachers who frequently requested the SLP's opinion about students' progress in the areas of language and literacy.

Audio Samples

The documentation of progress in therapy through audio samples was especially popular with students, teachers, and parents. These taped samples clearly documented students' progress, especially for oral reading and retelling tasks. The qualitative analysis of students' retellings over time often revealed improvements in specific aspects of communication such as a decreased incidence of behaviors associated with word retrieval difficulties or an increased use of article-noun agreement. When used with students with fluency difficulties, these tapes demonstrated students' ability to use fluency facilitating techniques, especially in more stress inducing situations such as speaking on the phone or with unknown adults. The audio samples also documented students' ability to speak with increased speech intelligibility while engaging in the same task over time, such as counting to ten or repeating a sentence that contained the student's targeted speech sounds.

Language Samples

The collection of language samples on an ongoing basis was an excellent means for the qualitative documentation of students' progress in therapy, especially when used with younger students or those with more significant speech or language disorders. In contrast to the more artificial language elicitation techniques associated with formal testing, these samples were collected in the context of the "real-time" competitive speaking environment of the group setting. These language samples were therefore more representative of students' improved communicative abilities in the school environment, especially when samples were collected in a variety of contexts such as the bilingual and ESL classrooms. For example, ongoing language samples collected on several preschoolers with mild-to-moderate delays and two kindergarten students with moderate-to-severe language disorders clearly documented their progress throughout out the year, not only in therapy, but in the classroom setting as well.

Therapy Notes

The systematic recording of observations of students' behaviors noted during the session or reported by teachers was one of the most effective means for the longitudinal documentation of progress in therapy. Because notes could not be written immediately after the therapy session, the implementation of a two-tiered system for data collection contributed to the documentation process. The brief notation written on the "post-it" notes facilitated in the recall of critical events so that a more complete recording of students' behaviors could be generated. In one case where notes could be

completed directly following the therapy session, the level of information contained in these notes was so extensive that quarterly reports were easily generated and a minimal level of testing was required for the annual review process.

These formative summaries of subjective and objective observations documented critical incidents that offered a means for analyzing ongoing patterns of behaviors that had the potential to impact students' progress in therapy and/or the classroom environments. For example, one student demonstrated a sudden and significant increase in his level of dysfluency over a four-week period that was also confirmed by his classroom and ESL teachers. A meeting was held with his parents to discuss this sudden regression in his communicative skills. Because the onset of these behaviors was carefully documented, his parents were able to identify factors related to this sudden change in behavior and a collaborative plan was devised to address his ability to communicate effectively within the classroom environment.

Interviews

Teacher's opinions of students' improvements in their communicative abilities in the classroom were usually elicited in informal interviews that were 5-10 minutes in duration. The quarterly reports focused on the students' progress in the therapy sessions, but also included comments about their communicative abilities within the classroom. On a number of occasions, classroom based therapy goals were planned collaboratively with teachers. In these cases, the specific information that documented the students' progress in the attainment of the therapy goals was explored in

interviews that were 10-15 minutes in duration. Quarterly reports for these students contained detailed descriptions of the students' progress in the classroom environment.

As part of the annual review process, parents were often interviewed about their child's progress in therapy as well as in the classroom. The information gained from these interactions with parents facilitated in the evaluation of the students' progress in school over the course of one year and helped to develop classroom-based therapy goals for the coming year.

Language and Literacy Rating Scales and Checklists

This study supported Silliman and Wilkinson's (1991) contention that when used in isolation, checklists do not provide details or insights into students' communicative abilities. However, when used in an open-ended format (e.g., how has the student demonstrated improvement in her active listening skills?), the CRLAI (Worcester Public Schools, 1999) clearly documented students' progress in the attainment of therapy goals. Because the CRLAI was based on the Massachusetts Curriculum Frameworks, it complied with the requirements of IDEA-97 that students' progress in the speech and language services be related to the general education curriculum.

Rubrics

The Student Writing Rubrics (Goddard School of Science and Technology, 1996) documented students' progress in the development of their writing skills. In

most instances, students' writing samples were the most easily discernable example of their progress in therapy. For example, several first grade students who also received inclusionary education services used pictures and attempted to record messages using letters to represent words at the beginning of the year. They then progressed to the use of favorite words written with both upper and lower case letters. At the end of the year, these students were beginning to write simple words and phrases using sound/symbol correspondence. Their progress in the regular education curriculum was clearly documented through their writing samples.

In summary, authentic assessments were found to be an ideal means for documenting students' progress in attaining educationally based communicative goals as required by IDEA-97. In contrast to the medical model of speech and language therapy that focuses exclusively on students' deficits, the authentic approaches to assessment used in this study provided insight into their progress within the general educational curriculum. The integration of intervention and assessment documented students' progress in their attainment of their therapy goals as outlined in their IEP.

CHAPTER 5

CONCLUSIONS

This study examined the evolution of assessment practices as a consequence of a SLP's involvement in educational change. This transformation took place at a challenging time. High-stakes assessments, anti-bilingual sentiment, and the movement for special education reform have focused on norm-referenced tests as a means of promoting high standards and measuring the efficacy of educational outcomes. As stipulated by IDEA-97, all students, including those with special needs and in the process of learning English, must be provided access to the least restrictive instructional program that is guided by the Curriculum Frameworks. As a result of this mandate, the majority of students with special needs are enrolled in regular education classes and receive inclusionary services. All students must participate in state-mandated testing including students with significant disabilities whose progress in the regular education curriculum must be measured through alternative forms of assessment. The disharmony created by the tension of these two reform movements has resulted in a high level of pressure for all educators who are being held accountable for the educational outcomes of all their students as measured by high-stakes assessments such as the MCAS.

The discord created by these two interrelated, but often times conflicting agents of change, has exerted a major impact for SLPs whose professional lives are greatly affected by the repercussions of these reform movements (Whitmire, 2000).

The purpose of this study was to explore the ways in which authentic measures of speech and language could be incorporated into the assessment protocol of a bilingual SLP with responsibilities for evaluating and treating a large caseload of Spanish-speaking students in an urban school system. The aim of this study was not to develop an "ideal model" but to investigate best practices under these work conditions. Authentic assessment procedures were found to be a viable alternative to standardized tests for differentiating language differences from language disorders, for selecting the language of instruction, and for documenting progress in therapy as mandated by IDEA-97.

IDEA-97 mandates a multi-stage process for determining eligibility for service. The first stage identifies the presence of a disability and in the case of the assessment of the communicative skills of a student from a culturally and linguistically diverse background, requires the differentiation of language disorders from language differences. The purpose of the second stage is to determine the extent to which the disability or condition negatively impacts the student's educational performance. The determination of eligibility is not directly related to the severity of the disability but to the adverse effect that the disability exerts on the student's educational performance. For example, a student with a significant degree of impairment as measured by standardized tests may not be considered disabled if classroom accommodations and modifications of the curriculum support the student's learning (ASHA, 2000).

According to IDEA-97, the determination of eligibility for service should not be restricted to the use of standardized assessments but should include a variety of

tools and strategies that provide functional and developmental information about a student's educational needs (ASHA, 2000). This study found that the pairing of authentic and standardized assessments was an effective means for the differentiation of language disorders from language differences. Given the multiplicity of factors with the potential to negatively impact results obtained from norm-referenced or authentic assessments used in isolation, both sources of information were needed to identify the presence of a communicative disorder. Because each source of data balanced and qualified the information contributed by the other, multiple sources of information were essential for the differentiation of language disorders from language differences. When used in conjunction with authentic measures such as interviews, acculturation scales, and curriculum-based language assessments, the judicious inclusion of norm-referenced measures was found to facilitate the determination of eligibility for service.

This study's multidimensional assessment model included a variety of tools and strategies that provided functional as well as developmental information about a student's educational needs as required by IDEA-97. In addition to the use of assessment devises associated with traditional speech-language evaluations, curriculum-based language assessments (such as miscue analyses, narrative assessments, story conferences, and writing rubrics) measured a student's linguistic abilities while performing integrative tasks aligned with the Curriculum Frameworks. These authentic assessment measures complied with the mandates of IDEA-97 by identifying the impact of the speech and language disorder on a student's ability to access the curriculum. Because of the flexibility of targeted tasks such as miscue

analyses, assessments could be targeted at a student's developmental level within the curriculum or supported classroom environment. This approach provided specific details about the student's ability to perform tasks at his/her "zone of proximal development" within the curriculum. When norm-referenced tests are used in a speech and language evaluation, a pupil's 's performance is compared to that of others of the same age or grade level, focuses on what the student does not know and contributes to feelings of failure. In contrast, this study's authentic approaches to assessment allowed a student to demonstrate what he/she knew at his/her developmental level and generated feelings of accomplishment.

According to IDEA-97, the identification of a disabling condition should also include a variety of alternative assessments such as interviews, targeted observations, reviews of work samples, and the completion of checklists (ASHA, 2000). Interviews with parents, students, and teacher were used to identify the specific contexts within the curriculum that were considered to be problematic for the student. Observations and the examination of work samples were integral components of curriculum-based language assessments and served as a means for understanding the interaction of the demands of the classroom environment and the student's ability to respond effectively to these demands (ASHA, 2000; Nelson, 1998). Although behavioral checklists and rating scales were occasionally incorporated into the process of differentiating language differences from language disorders, they were found to be most productive when used as structured interviewing devices.

IDEA-97 stipulates that evaluation procedures must be conducted in the native language of the student and must not discriminate against the individual because of

racial or cultural factors. Because one of the major purposes of this study was to determine the language of instruction for bilingual students with communicative disorders, the assessment of language dominance and proficiency in both languages was integrated into all aspects of the evaluation process. These procedures included record reviews; interviews with parents, students, and teachers; and the analysis of a student's level of acculturation. Authentic assessment procedures such as miscue analyses and narrative language sampling were also used to evaluate the level of proficiency of the student's cognitive/academic language skills in both languages.

IDEA-97 recommends that an ethnographic approach be implemented for the gathering of information about the influence of cultural factors on a student's communicative abilities (ASHA, 2000). In this study, the use of ethnographic interviews to identify "zones of significance" combined with the analysis of a student's level of acculturation facilitated in the differentiation of a language difference from a disorder. That is, the applicability of the results of standardized assessments conducted by evaluation team members was interpreted within the context of a student's level of acculturation. The influence of factors such as language loss and second language acquisition on a student's dual language abilities was determined through the analysis of his/her level of acculturation. Thus the analysis of a student's level of acculturation was a vital source of information for differentiating language differences from language disorders.

In addition, the results from this study supported the contention that the inclusion of both standardized and authentic measures decreased the likelihood of assessment bias. As previously indicated, this study found that the implementation of

a multidimensional assessment model that incorporated interviews, acculturation scales, curriculum-based language assessments, and dynamic assessments with the judicious use of standardized tests facilitated in the determination of eligibility of services. Because the results from the multiple sources of data were triangulated and patterns of linguistic behaviors were analyzed within a variety of contexts, the validity and reliability of the outcomes were enhanced and the level of assessment bias was decreased.

One of the presumed advantages of norm-referenced tests is their provision of a valid and reliable estimate of an individual's linguistic abilities (Damico, 1991; Duran, 1989; McCauley & Swisher, 1984). A major criticism of authentic assessment procedures is related to the subjectivity of the judgements made about a student's performance; however, these arguments are based upon traditional perspectives of validity and reliability (Leslie & Dussault, 1997).

Test validity refers to the "appropriateness, meaningfulness, and usefulness of specific inferences made from test scores....The inferences regarding the specific uses of the test are validated, not the test itself" (American Psychological Association, 1985, p. 9). This study's pairing of qualitative and quantitative assessments increased the likelihood of the appropriate interpretation of test scores from norm-referenced tests. In addition, the addition of norm-referenced tests to the authentic assessment protocol increased the probability that valid inferences would be generated about a student's performance with the classroom curriculum. Thus the information gained from both sources of data resulted in appropriate and useful conclusions about a student's communicative abilities within meaningful contexts as required by IDEA-97.

Messick defined test validity as the reasonableness of the conclusions drawn from the assessment data and the reasonableness of the action taken based upon that information (1989). The collection of data from multiple qualitative and quantitative data sources increased the reasonableness and accuracy of the information obtained in this study. The findings from this study were in agreement with Leslie and Dussault (1997) who argued that because authentic assessment data are used to improve a student's performance within the classroom setting, this information has higher levels of validity than the results from norm-referenced tests. As previously stated, results from norm-referenced tests cannot be used to develop specific goals for intervention (Merrell & Plante, 1997; Plante & Vance, 1994). In initial speech-language evaluations, results from authentic assessments qualified and clarified the results of norm-referenced tests. The information obtained from the authentic assessments led directly to the establishment of goals for speech and language services. In cases where students were enrolled in therapy, the ongoing observation of a student's progress in therapy guided the subsequent instruction. Assessment was a daily activity that analyzed, synthesized, and documented information from multiple sources. Based upon these assessment data, reasonable actions were taken and the validity of the assessment procedures was increased.

Leslie and Dussault (1997) further noted that the validity of the authentic assessment data is increased as a result of the frequent observations of the targeted behaviors within a variety of contexts that leads to more accurate inferences made about the targeted behaviors. This study's multidimensional assessment model collected and triangulated data from multiple informants and within a variety of

contexts. As a result, the incorporation of multiple quantitative and qualitative sources of information increased the reasonableness of the conclusions made from the data and validity of the conclusions drawn from the data.

Reliability refers to the dependability of a test as evidenced in the consistency of scores upon repeated measurement of the same group (Nelson, 1998). One way to establish the reliability of an assessment procedure is to examine the consistency of the observation of targeted behaviors within certain contexts (Leslie & Dussault, 1997). In this study, test-retest reliability was ascertained through repeated administrations of the same activity (e.g., miscue analyses using texts of the same level of difficulty) or of the same activity in different contexts (e.g., narrative samples after reading a story and while viewing a wordless book).

According to Leslie and Dussault (1997), interscorer reliability can be established by determining the level of agreement arrived at by independent observers of the same behaviors. These authors postulate the following question as a check of the interscorer reliability of an assessment procedure: If two examiners were to observe a student, would they draw the same conclusion about the presented behaviors? Two ways in which interscorer reliability was obtained in this study were through the triangulation of data from interviews with the student, parents, and teachers and the comparison of results from miscue analyses with those obtained from the teacher or school psychologist. These practices enhanced the reliability of the conclusions that were drawn about a student's communicative abilities.

Finally, authentic assessment procedures not only differentiated language differences from language disorders, but also identified specific learning strategies and

classroom accommodations that facilitated the student's ability to access the curriculum. Many of the interviewed educators expressed concern about their lack of training in language facilitation techniques required by their students with specialized learning needs. Specific strategies or scaffolding techniques were identified through curriculum-based learning tasks that were administered in a dynamic manner, thus contributing to positive outcomes of learning. The completion of authentic assessment procedures identified ways in which language learning could be supported by classroom teachers, thereby increasing the accessibility of the general education curriculum as required by IDEA-97.

The inclusion of authentic measures that targeted the assessment of language dominance and proficiency was found to be an essential component in determining the language of instruction for bilingual students with communicative disorders. The assessment of language dominance investigated the language that the student used in the majority of contexts as well as the language used most often when either language would be appropriate. In addition, the authentic assessments implemented in this study identified the language whose rules or structures seemed to influence the production of the other language most frequently (Ortiz & Garcia, 1990). Finally, language proficiency or a student's skill level in each language was evaluated using a dual language multidimensional assessment model that included procedures such as miscue analyses, the examination of writing samples, and the determination of level of acculturation.

In general, standardized measures of language proficiency focus on basic interpersonal communication skills, rather than the cognitive/academic language skills

required for success in the monolingual program. Given the pervasive influence of high-stakes assessments on all aspects of education combined with the poor performance of Latino children on these tests (Massachusetts Department of Education, 2000), the study found the evaluation of cognitive/academic language skills to be imperative, especially when used to determine the language of instruction for these students. Authentic assessment techniques, such as miscue analyses, the examination of classroom artifacts, and narrative samples, helped to ascertain whether or not students presented with the requisite English proficiency skills to meet the cognitive and academic demands of the monolingual program. The corroboration of these curriculum-based assessments with information derived from acculturation scales, interviews, and observations provided a holistic perspective on a student's ability to access the general education curriculum in English. The utilization of a dual language multifaceted curriculum-based assessment model facilitated in the selection of the language of instruction for bilingual students with communication disorders.

Authentic assessments were found to be an ideal means for documenting a student's progress in attaining specific educationally-based communication goals as required by IDEA-97. The use of ongoing curriculum-based language assessments and ethnographic interviews with the student as well as teachers and parents as documented in detailed therapy notes, served as a means for ascertaining the communicative abilities of a student in relation to the curriculum. In contrast to the more traditional therapeutic model that focuses on a student's weaknesses, the careful documentation of his/her ability to meet the linguistic demands of the educational program provided a holistic perspective of the student's strengths and achievements

across all aspects of the curriculum. The ongoing integration of intervention and assessment assisted in the planning of therapeutic activities geared to the learning needs of the student within the context of the curriculum and decreased the reliance on more formalized testing procedures when determining eligibility for continued services.

This study has found that a dual language, multidimensional, and balanced approach to assessment, which incorporated both standardized and authentic assessment procedures, was required for differentiating language differences from language disorders, for selecting the language of instruction, and for documenting progress in therapy. However, the ramifications of the mandates of IDEA-97 have exerted an impact on school-based SLPs who are seeking a balanced approach to their changing roles as evaluators and interventionists. As SLPs face the new millennium in the shadow of IDEA-97, the inclusion of curriculum-based assessments and intervention is no longer an option but a requirement. The resultant paradigm shift in the definition of the role of the school-based SLP will demand that they incorporate authentic and curriculum-based assessments and collaborative classroom-based practices (Whitmire, 2000). They will need to explore both individually and collectively, how they as clinicians can contribute to the success of communicatively challenged students in the general education curriculum.

Future research is needed to answer several questions. How can school-based SLPs with large caseloads incorporate authentic and curriculum-based assessments into their already busy schedules? How can these assessments be completed within the time frames mandated by IDEA-97 without canceling multiple therapy sessions?

What is the role of SLPs in preparing students with communication disorders for high-stakes assessments such as the MCAS? In helping students to prepare for these tests, how can SLPs integrate their clinical knowledge with classroom-based practices without becoming tutors or teaching assistants? How can SLPs contribute to the development of alternative assessments for students with significant special needs?

The ability of SLPs to redefine their role in the context of educational accountability, high standards, and measurable outcomes is quickly entering the critical stage where the granting of a high school diploma for the majority of their students with special needs will be determined by the students' score on a high-stakes test.

APPENDICES

APPENDIX A

PARTICIPANT INTERVIEW QUESTIONS

Students

1. What do/did you like best about the portfolio and/or portfolio conference?

   Probe: How do you feel when you look at the work in your portfolio? Why?

   Probe: How do you feel when other people (e.g., the speech-language pathologist or your friends) look at the work in your portfolio? Why?

   Probe: How did you feel when your mother or teacher looked at the work in your portfolio? When did you feel the happiest? The proudest? Why?

2. How does/did the portfolio and/or portfolio conference show what you have learned this year?

   Probe: What does the work in your portfolio tell other people (e.g., your mother, teacher, friends) about you?

   Probe: What is your favorite piece of work? How does it show others what you have learned this year?

   Probe: (Speech-language pathologist chooses an artifact from early in the year and a more recent piece of related work e.g. literature responses or self-portraits, How do these two pieces show us what you have learned this year?

What can we change about our portfolios to make them better?

Probe: If you were to tell a friend the best thing about portfolios, what would you say? If you were to tell a friend what we need to do to make them GREAT, what would you say?

Probe: If the principal gave us lots of money to make our portfolios better, what could we do?

Parents

1. What did you like best about the portfolio and/or portfolio conference? Why?

   Probe: What did you like best about your child's collection of work? Why?

   Probe: Which piece from your child's collection of work did you like best? Why?

   Probe: What did you like best about the meeting with your child and the speech-language pathologist? Why?

2. How did the portfolio and/or portfolio conference demonstrate your child's progress this year?

   Probe: How did your child's collection of work demonstrate what he/she has learned this year?

   Probe: Which piece of work most clearly showed your child's progress this year? How did it show this progress?

Probe: Do you feel that you now have a better understanding of how your child is doing in school and/or speech therapy? Why or why not?

3. Did you find portfolios easier or more difficult to understand than testing techniques that have been used in the past? Why or why not?

Probe: Was this collection of work easier or more difficult to understand that the testing that has been done in the past? Why or why not?

Probe: Now that you have seen a collection of your child's schoolwork, do you think that it will be easier or harder to talk him/her about how she/he is doing in school? Why or why not?

Teachers

1. What are the advantages of authentic/portfolio assessments? What are the disadvantages of authentic/portfolio assessments?
2. Based upon your experiences, how do your students and their parents view authentic/portfolio assessments? How do administrators view these assessment practices?
3. How do you adapt authentic/portfolio assessment practices when working with students with special needs or students from culturally and linguistically diverse backgrounds?

4. How can speech-language pathologists better reflect students' classroom performance in their assessments?
5. What effect will high-stakes assessments have on authentic/portfolio assessment practices? What is the future for authentic/portfolio assessment practices?

APPENDIX B

INDIVIDUAL STUDENT CONFERENCE FORM

(Adapted from Success For All Foundation, 1998)

Story: A Snowy Day by E. J. Keats

1. What was the name of the story?

2. Who were the story characters?

4. What did Peter see when he looked out the window?

5. What did Peter do in the snow?

6. Why didn't he play with the older children?

7. What do you think Peter thought about during his bath?

8. What happened to the snowball that he put in his coat pocket?

9. What was your favorite part of the story? Why?

APPENDIX C

RUBRIC FOR INDIVIDUAL STUDENT CONFERENCES

(Adapted from Porch, 1967)

4 = Accurate, complete, prompt response

3 = Minimal cue- Accurate, but delayed response; one repetition of question

2 = Moderate cue- Partial response; significantly delayed response, or two repetitions of question

1 = Maximal cue- Inaccurate, but related, very significantly delayed response, or three repetitions of question

0 = No response, response unrelated to the question, or four or more repetitions

APPENDIX D

LEVELED READING BOOKS

Spanish

Level	Title	Author/Series	Publisher/Distributor
Kindergarten/Grade One			
A	Mi cuarto	Libros Carrusel	Dominie Press
B	El oso	Libros Carrusel	Dominie Press
Grade One			
C	La mochila de Lin	Lester, H.	Scott Foresman
	Beto y Pedro	Franco, B.	Scholastic
D	Primer día de escuela	Libros Carrusel	Dominie Press
F	El paseo de Rosita	Hutchins, P.	Scholastic
Grade One (late)			
I	¿Eres tu mi mama?	Eastman, P. D.	Random House
Grade Two (early)			
J	Rosaura en bicicleta	Barbot, D.	Ediciones Ekare
Grade Two			
K	Alexander, que era rico el domingo pasado	Viorst, J.	MacMillan

L	Arturo y sus problemas con el profesor	Brown, M.	Little, Brown, & Co.

Grade Three

N	Jorge el curioso	Ray, H. A	Scholastic
	Jorge el curioso en el hospital	Ray, M. & H. A.	Scholastic
O	El sancocho del sábado	Torres, L	Scholastic

Grade Four

Q	El secreto de la llama	Palacios, A.	Troll
R	El misterio de La Isla de las Especias	Keens-Douglas, R.	Annick Press

Grade Five

S	Leyendas puertorriqueñas	Rodriguez , A. & Puigdollers, C.	Northeast Center for Curriculum Development
	La llorona	Anaya, R.	Houghton Mifflin

English

Level	Title	Author/Series	Publisher/Distributor

Kindergarten/Grade One

A	I Can See	Carousel Earlybirds	Dominie Press
B	Cat on the Mat	Wildsmith, B.	Oxford

Grade One

C	I Went Walking	Williams, S.	Harcourt Brace
D	My Messy Room	Hello Reader	Scholastic
F	Itchy Itchy Chicken Pox	Hello Reader	Scholastic

Grade One (late)

I	Noisy Nora	Wells, R.	Scholastic

Grade Two (early)

J	Danny and the Dinosaur	Hoff, S.	Scholastic

Grade Two

K	Arthur's Honey Bear	Hoban, L.	Harper Collins/General
L	Miss Nelson Is Missing	Allard, H.	Houghton Mifflin

Grade Three

N	The Enormous Alligator	Dahl, R.	Puffin Books
	Curious George Rides a Bike	Rey, M.	Scholastic
O	Socks	Cleary, B.	Avon Books

Grade Four

Q	School's Out	Hurwitz, J.	Scholastic
R	Old Ben Bailey	Justus, M.	Celebration Press

	Meets His Match		

Grade Five

S	Danger in the Deep	Coombs, C.	Celebration Press
	The People Could Fly	Hamilton, V.	Knopf

APPENDIX E

MISCUE TRANSCRIPTION SYSTEM

(Adapted from Clay, 1993; Goodman, Watson, & Burke, 1987)

Substitutions	The man run ran
Repetition or re-read	She said that she was going
Re-read more than once	She said that R3 she was going
Self-correction	The man run ran (c)
Omissions	The man ran very fast
Insertions	The man ran very fast
Dialect and other language variations	The girl jumped ^ jump (d) over
Intonation shift	They will record
Non-word	It had two equal $ sidse sides
Misarticulations	It had two equal θ/s (m) sides
Assistance from the examiner	His [tumb] [thumb] started to bleed.

APPENDIX F

LANGUAGE AND READING OBSERVATION GUIDE

(Adapted from Beaver, 1997; Burke, 1980; Clay, 1993; Weaver, 1980, 1994)

NAME _____ DATE _____

GRADE _____READING LEVEL _____ LANGUAGE: _____

NAME OF TEXT: _____

COMMENTS: _____

_____

INTRODUCTION:

Read the title and look at pictures in story. Ask student to predict what is going to happen in the story.

As student viewed pictures, he/she evidenced: emerging adequate ability to use pictures

treated each picture as a separate event emerging story logical story

2nd grade and above: What do you think might happen in the story?

Predictions: unlikely logical multiple

ORAL READING:

READ: syllable-by-syllable word-by-word in short phrases in longer phrases

OBSERVED: punctuation

REREAD FOR: phrasing punctuation stress/intonation

INTONATION: emerging developing generally effective

READING RATE: slow inconsistent adequate too rapid adjusted as needed

CONSTRUCTED MEANING THROUGH: pictures rereading letters/sounds syllables

multiple attempts pausing meaning syntax

self-corrected: never rarely sometimes often

no observable behaviors

REQUESTED HELP: often sometimes rarely never

MISCUE ANALYSIS:

Were miscues due to a language difference?	Never	Rarely	Sometimes	Often
Did miscues go with preceding text?	Never	Rarely	Sometimes	Often
Did miscues go with following text?	Never	Rarely	Sometimes	Often
Did miscues preserve essential meaning?	Never	Rarely	Sometimes	Often
Were miscues self-corrected?	Never	Rarely	Sometimes	Often
Were miscues meaning preserving or self-corrected?	Never	Rarely	Sometimes	Often

STORY RETELL:

Unaided:

Aided:

Students in 3rd grade and above: In one or two sentences, tell me what the story was about:

Initial Retelling Included:	main characters	setting	important details	
	vocabulary/phrases from story		starts retelling at beginning	
	events in sequence		events out of sequence	
	ending		provides information not in story	
With Question Prompts:	main characters	setting	important details	
	vocabulary/phrases from story		starts retelling at beginning	
	events in sequence		events out of sequence	
	ending		provides information not in story	
Level of Cueing:	Mild	Moderate	Maximal	

Did student evidence improved comprehension when story reread silently?

Level of Improvement:	Significant	Moderate	Mild	No difference
:	poor	needs improvement	fair	satisfactory
Comprehension/Recall of Characters/Events	1	2	3	4
Comprehension of Inferential Information	1	2	3	4
Organization of Narrative	1	2	3	4
Content of Narrative	1	2	3	4
Vocabulary	1	2	3	4
Sentence Structure	1	2	3	4

FOLLOW-UP QUESTIONS:

Did you like this story? Why or why not?

What did this story make you think of?

READING INTERVIEW:

When you are reading, what do you do when you come to something that you do not know?

What else do you do?

Who is a good reader that you know? What makes that person a good reader?

Are you a good reader? Why or why not?

Are you a better reader in English or Spanish?

Why is reading easier in _____ than _____?

What makes reading in _____ so hard?

Do you prefer to read in English or Spanish? Why?

Level of cueing for questions: Mild Moderate Maximal

APPENDIX G

READING COMPREHENSION RUBRIC

(Adapted from Beaver, 1997)

Early Readers (Kindergarten- Middle First Grades)

Retelling reflects very little understanding:

- unorganized
- important details missing
- incorrect information
- misinterpretation

Early Transitional Readers (Late First- Middle Second Grades)

Retelling reflects some understanding:

- somewhat organized
- focuses on parts rather than wholes
- events out of sequence
- includes some details about characters and events
- some misinterpretations
- links to personal experiences

Late Transitional Readers (Middle Second- Early Third Grades)

Retelling reflects an adequate level of understanding:

- organized but may be choppy
- generally sequenced appropriately
- includes main ideas, details about characters, setting, and events
- literal interpretation
- links to media or other events

Extending Readers (Third- Fifth Grades)

Retelling reflects very good understanding:

- adequately organized
- adequately sequenced
- includes main idea, important details about characters, setting, and events
- reveals use of background knowledge and experiences to interpret story
- uses vocabulary and special phrases from the story
- links to other literature

APPENDIX H

RATING SCALE-NARRATIVE PRODUCTION AND NARRATIVE SKILLS CHECKLIST

(Adapted in collaboration with L. Ohl from Singer & Bashir, 1998; Worcester Public Schools, 1995)

Rating Scale-Narrative Production

Score	Characteristics
4	The student's narrative is **satisfactory/good**. The narrative … • focuses on the topic and includes an adequate level of supporting ideas, details, or examples (C) • contains an organizational pattern with a limited number of deviations noted (O) • demonstrates a sense of completeness or wholeness (C/O) • evidences adequate word choice, but is lacking in precision on rare occasion (V) • conveys complete thoughts in general (S) • evidences isolated morphosyntactic errors, however, the content is not compromised (S) • contains compound and complex sentences that are generally well-formed (S)
3	The student's narrative is **fair**. The student … • focuses on the topic in general, however, it contains some extraneous or loosely related information (O) • evidences an organizational pattern, however, occasional deviations are noted. A limited number of self-corrections are demonstrated.(O) • demonstrates a sense of completeness or wholeness (O) • provides some supporting ideas with details or examples, however, other ideas may not be as well developed (C) • evidences adequate word choice, however, it may be lacking in precision on occasion (V) • contains a variety of sentences structures, however, many are simple constructions (S) • contains a limited number of morphosyntactic errors that impact the comprehensibility of the narrative on occasion (S)

2	The student's response **needs improvement**. The student … • relates minimally to the topic and/or offers a limited amount of relevant information with few supporting details or examples (C) • attempts an organizational pattern that contains some transitional devices, however, deviations occur. Isolated self-corrections are noted. (O) • lacks a sense of completeness or wholeness (C/O) • contains isolated supporting ideas, details, or examples (C) • evidences adequate word choice, but it is limited, predictable, and sometimes vague (V) • uses a variety of sentence structures, however, the majority are simple constructions (S) • contains consistent morphosyntactic errors that impact the comprehensibility of the narrative (S)
1	The student's response is **poor**. The student … • relates minimally to the topic and demonstrates very limited development of supporting ideas, details, or examples (C) • evidences limited development of an organizational pattern and contains frequent deviations from the topic (O) • develops supporting ideas inadequately or illogically (C/O) • evidences limited word choices that negatively impact comprehensibility (V) • contains simple sentence structures and multiple morphosyntactic errors (S)

KEY

(C) = content
(O) = organization
(C/O) = content/organization
(V) = vocabulary
(S) = structure

Narrative Skills Checklist

(Adapted from Westby, Van Dongen, & Maggart, 1989)

Process/Content

Preschool

Ability to label _____

Awareness of animate/inanimate:
animate beings act & inanimate
acted on _____

Awareness of physical cause-effect _____
Beginning awareness of linear time _____

Early Elementary

Awareness of psychological causality
for primary emotions _____
Awareness of situations cause emotions _____
Theory of mind (people think/feel) that
allow for perspective taking _____
Ability to conceptualize near future _____
Scriptal knowledge of common char.
(wolves are bad) _____

Further development of psychological
causality:
secondary or cognitive emotions _____
Awareness of char attributes w/ story
elements of setting/events- can compre.
& predict novel behavior of character _____
Understanding longer time frames _____
Meta-awareness- need to plan _____
How to plan _____
Need to justify plan _____

Structure

Isolated labeling: description of objects _____
actions _____ characters _____
surroundings _____
No interrelationships among elements _____

Action sequence, not causally related _____
May be chronologically ordered _____

Characters act independently of each other _____
May have central character or theme _____

Set of actions/events that cause events_____
No planning involved_____

Abbreviated episode_____
Centering & chaining present _____
Characters engage in cause-effect actions_____
Describes goals/intentions of characters _____
Planning must be inferred _____
Components: Initiating event _____
Response _____
Consequence _____

Complete episode

Centering/chaining present _____
Describes goals of characters _____
Describes intentions of characters _____
Some evidence of planning _____
Components: Initiating event (problem) _____
Characters' reaction to problem _____
Plan _____
Attempt (carrying out plan) _____
Consequence _____

Late Elementary School

Ability to see char change/growth _____
Ability to perceive deception/trickery _____
Ability to deceive or trick _____
Awareness of time cycles (seasons/years) _____
Beginning awareness multi-meaning words _____
figurative vs literal meanings _____

Elaborated Stories:
Complex episode: single episode story w/ multiple plans, attempts, or consequences _____
Interactive: story told from point of view of more than one char. _____
Multiple episodes: story has more than one chapter w/ each chapter having story grammar elements- initiating event/ response/plans/attempts/consequence. _____
Early stages- sequential _____
Later stages- may be embedded _____

Adolescent/Adult

Ability to engage in metanarrative discussion:
Narrative structure _____
Interpretation of characterization _____
Themes _____
Plots _____
Understanding of abstraction of time/space _____
Ability to understand flashbacks_____
Ability to understand allegories _____

Metaphoric stories: Usual structure may be modified in novel ways for humorous or metaphoric purposes. _____

KEY:

S= Spontaneous C= Could not elicit

E= Elicited NA= Not applicable/Not assessed

APPENDIX I

INTERVIEWS

Student Interview

(Adapted from Nelson, 1992)

1. Student's description of what is hardest about school.

2. Student's description of what is best about school.

3. Student compares and contrasts present level of academic achievement to his/her performance level in bilingual program (in cases where student is presently enrolled in a monolingual program, but had previously attended a bilingual program).

4. Student's prioritized list of changes to be made.

5. Anecdotal evidence of recent classroom events that made the student feel really bad.

6. Student's ideas about the future.

Teacher Interview

(Adapted from Nelson, 1992)

1. Objective information about the student's academic performance- formal tests and classroom levels of performance. Teacher compares and contrasts student's performance to other students who have previously attended bilingual classes or who have had similar linguistic and cultural experiences.

2. Student's strengths.

3. A prioritized review of problems the teacher identifies as most important.

4. Anecdotal description of recent classroom events in which the student experienced difficulty.

5. Description of aspects of curriculum that present the greatest difficulty to the student and are of the most concern for the teacher.

6. Teacher's view of the student's potential within the current school year and in the future.

Parent Interview

(Adapted from Gonzales & Kayser, 1997; Nelson, 1992; Restrepo, 1998)

1. Early development- Did they suspect a problem early?

2. Medical history-especially middle ear infections

3. Family history of language/learning disabilities or special needs. Comparison of child's language and learning abilities to peers with similar cultural and linguistic experiences.

4. Educational history
   When did problems first show up in school?

   Did decoding problems show up early, or did the problems show up in 3rd or 4th grade when it became more important to read longer texts for meaning?

5. Anecdotal evidence of specific problems within the past year or so.

6. A prioritized view of the problems the parents view as most critical.

7. Parent's goals for the child's future.

Listening/Understanding/Remembering Interview

(Adapted from Burke, 1980)

1. What do you do when your teacher tells you to do something, but you were not listening so you don't know what to do?
   Do you ever do anything else?
   What do you do when your teacher says or tells you something that you do not understand?
   Do you ever do something else?

What do you do when you do not remember what you have to do in class?
Do you ever do something else?

2. What sorts of things help you to listen in school?

3. What sorts of things help you to understand things in school?

4. What sorts of things help you to remember things in school?
5. What do you think you could do to listen better in school?

6. What could you do to understand things better in school?

7. What could you do to remember things better in school?

8. Do you remember things better in English or Spanish?

9. Why is it harder to remember things in _____ than _____?

10. Why is it easier to remember things in _____ than _____

Word Retrieval Interview

(Adapted from Burke, 1980)

1. When you are talking and you can't think of the word that you want to say, what do you do?
   Do you ever do anything else?

2. Who is someone that you know that speaks very well?

3. What makes them a good speaker?

4. Do you think that he or she ever has a hard time thinking of the word(s) that he/she wants to say? What do you think that he/she might do so that other people understand what he/she is trying to say?

5. If you knew someone was having a hard time thinking of the words that he/she wanted to say, what would you tell that person to do?

6. What sorts of things should you do so that people understand what you are trying to say when you can't think of the word that you want to say?

7. Do you remember words better in Spanish or English?

8. Why is it easier to remember words in _____?

9. Why is it harder to remember words in _____?

APPENDIX J

ORAL READING OBSERVATION GUIDE

(Adapted from Fountas & Pinnell, 1996)

	Spanish			English		
	beginning	developing	well-developed	beginning	developing	well-developed
Levels A & B (K/1)						
Handles books- front to back, turning pages						
Controls L to R, return sweep						
Notices & interprets details in pictures						
Uses oral language relative to text						
Matches word by word (precise pointing)						
Pays close attn. to print- notices some features of letters/words						
Locates familiar & new words						
Remembers & uses language patterns						
Uses knowledge of syntax as source of information						
Uses oral lang. in combo w/ pointing-matches voice with words						
Predicts what makes sense						
Self-monitors- Checks reading by word-by-word matching, notices mismatches in meaning or lang.						
Level C (1st)						
Uses visual info. to help predict, check, & confirm						
Controls word-by-word matching of voice with print						
Uses pictures to predict meaning as well as particular words						
Predicts from events what will happen next						
Checks pictures with print						
Uses known words as anchors						
Moves fluently through text while reading for meaning						
Solves some unfamiliar words independently						
Engages independently in behaviors listed in Level B						
Level D						
Controls early strategies (word-by-word matching/directional movement)						
Moves away from finger pointing as eyes take over process						
Uses pattern and syntax to read with phrasing						
Checks reading using letter-sound relationships, words, & parts of words						
Rereads to confirm reading or problem solve						

	Spanish			English		
	beginning	developing	well-developed	beginning	developing	well-developed
Checks one source of information against another to confirm, make another attempt, or self correct						
Actively reads for meaning						
Level E						
Tracks print with eyes except at points of difficulty or on novel texts						
Uses knowledge of syntax & meaning to read with phrasing						
Rereads fluently						
Solves new words while maintains focus on meaning						
Rereads to check, confirm, & search						
Cross-checks one source of information with another						
Self-corrects using multiple sources of information						
Predicts what will happen next & rereads to confirm						
Uses known words to get to words not known yet						
Relates one text to another						
Uses more info from print to construct meaning of story						
Level F						
Shows awareness of punctuation- uses for phrasing & meaning						
Searches visual information to figure out new words while reading						
Uses syntax of written lang. to predict, then checks accuracy of prediction						
Analyzes new words/checks against what makes sense or sounds right						
Controls early strategies even on novel texts						
Reads with fluent phrasing & attention to meaning						
Moves quickly through text						
Uses known words & parts of words & letter-sound relationships to get to new words & checks against other information such as meaning						
Uses multiple sources of information to self-correct						
Levels G & H						
Moves thru text using pic/print in integrated way & attends to meaning						
Solves new words using word analysis, checks against meaning						
Self-corrects close to point of error						

	Spanish			English		
	beginning	developing	well-developed	beginning	developing	well-developed
Rereads to check & search						
Discusses ideas from story- indicates understanding						
Discusses characters- indicates understanding & interpretation						
Manages variety of texts- fiction & informational						
Connects text to other texts						
Level I (Late 1st)						
Shows fluent & phrased reading- specially when rereading						
Problem solves new words on initial reading						
Checks reading against meaning						
Uses info. sources (meaning, syntax, & visual) in integrated way while focusing on meaning						
Makes connection between texts thru discussion, art, or writing						
Shows understanding of & empathy w/ character (art/discussion/writing)						
Moves toward easy, fluent reading even for unfamiliar & difficult text. Shows less overt problem solving						
Self-corrects at point of error- fewer returns to beginning of sentence/phrase						
Copes with unfamiliar concepts						
Gains momentum while moving thru text- knowledge of how text works & what likely to say						
Level J (Early 2nd)						
Uses skills & strategies effectively on variety of texts						
Shows understanding of story or text through discussion/art/writing						
Moves flexibly from fiction to nonfiction						
Uses ideas from reading in writing						
Summarizes or extends given text						
Level K (Late 2nd)						
Uses multiple sources of information in integrated way						
Reads silently much of time						
Effectively & efficiently analyzes words						
Uses variety of word analysis strategies without losing meaning/fluency						

	Spanish			English		
	beginning	developing	well-developed	beginning	developing	well-developed
Reads in phrased fluent way over longer stretches of text						
Shows (writing/discussion/other media) that understands/interprets stories from different perspectives & empathizes w/ characters						
Uses text structure (narrative & logic) to predict likely sequence of events or to analyze & critique text						
Sustains plot & characters over several days						
Level L						
Same behaviors for Level K, but exhibited in						
connection w/ longer stretches of text						
More difficult vocabulary, ideas, language structures						
More complex ideas & topics						
A greater range of genres						
Level M						
Same behaviors for K & L Also:						
Uses texts as references						
Searches for & finds information in texts						
Interprets texts from a variety of perspectives						
Reads critically						
Shows understanding of subtleties of plot & humor						
Reflects on personal response in relation to						
how others see the text						

APPENDIX K

BACKGROUND INFORMATION

(Adapted from Acevedo, Hardee, & Fernandez, 1997; Gonzales & Kayser, 1997; Ortiz & Garcia 1990)

1. Medical History:
2. Speech-Language Development:
3. Family History of Special Needs:
4. Student's Place of Birth:

   Mother's Place of Birth:

   Father's Place of Birth:

   Grandparents' Place(s) of Birth:
5. Educational Level of Parents

   Mother Location

   Father Location
6. Current Employment

   Mother Father
7. Date of arrival to U.S./Mainland:
8. How and when English was first introduced to student:
9. Educational placement of siblings:
10. Age entered school:

Educational Background:

				Language of Classroom					
	Location	Program	Span.	1	2	3	4	5	Eng.
Pre-school (3 years)				1	2	3	4	5	
Pre-school (4 years)				1	2	3	4	5	
K				1	2	3	4	5	
1				1	2	3	4	5	
2				1	2	3	4	5	
3				1	2	3	4	5	
4				1	2	3	4	5	
5				1	2	3	4	5	
6				1	2	3	4	5	
Middle School									
High School									

Frequent Absences:

Frequent Moves:

Retentions:

REFERENCES

Acevedo, M. A., Hardee, W. P., & Fernandez, M. R. (1997, November). Finding solutions for service delivery issues in LEP students. Paper presented at the American Speech-Language-Hearing Association National Conference, Boston, MA.

Adler, S. (1981). Poverty children and their language: Implications for teaching and treating. New York: Grune and Stratton.

Allington, R., & Cunningham, P. (1996). Using running records. In I. C. Fountas & G. S. Pinnell. Guided reading: Good first teaching for all students (pp. 89-96). Portsmouth, NH: Heinemann.

American Psychological Association. (1985). Standards for educational and psychological testing. Washington, DC: Author.

American Speech-Language-Hearing Association. (1991, May). Our multicultural agenda 2000. ASHA, 38-39.

American Speech-Language-Hearing Association. (1998, June). Guidelines on the roles and responsibilities of the school-based speech-language pathologist-draft copy. Rockville, MD: American Speech-Language-Hearing Association Ad Hoc Committee on the Roles and Responsibilities of the School-Based Speech-Language Pathologist.

American Speech-Language-Hearing Association (2000). IDEA and your caseload: A template for eligibility and dismissal criteria for students ages 3 to 21. Rockville, MD: Author. [on-line] Available: http://www.asha.org./idea/idea_caseload.htm

American Speech-Language-Hearing Association (in press). Roles and responsibilities of speech-language pathologists with respect to reading and writing in children and adolescents: Position statement, guidelines, and technical report. Rockville, MD: Author.

Anderson, J. L. (1989). Acculturation, adaptive behavior performance, and school placement of children who are culturally and linguistically different. Unpublished doctoral dissertation, University of Minnesota.

Anderson, R. T. (1995). Spanish morphology and syntactical development. In H. Kayser (Ed.). Bilingual speech-language pathology: A Hispanic focus (pp. 41-74). San Diego: Singular Publishing Company.

Anderson, R. T. (1996). Assessing the grammar of Spanish-speaking children: A comparison of two procedures. Language, Speech, and Hearing Services in Schools, 27(4), 333-344.

Apel, K. (1999). An introduction to assessment and intervention with older students with language-learning impairments: Bridge from research to clinical practice. Language, Speech, and Hearing Services in Schools, 30(3), 228-230.

Barona, A., & Miller, J. A. (1994). Short Acculturation Scale for Hispanic Youth (SASH-Y): A preliminary report. Hispanic Journal of Behavioral Sciences, 16(2), 155-162.

Bartoli, J., & Botel, M. (1988). Reading/learning disabilities: An ecological approach. New York: Teachers College Press.

Beaver, J. (1997). Developmental reading assessment. Glenview, IL: Celebration Press.

Berry, J. W. (1980). Acculturation as varieties of adaptation. In A. Padilla (Ed.), Acculturation: Theory, models, and some new findings (pp. 9-25). Boulder, CO: Westview Press.

Bissex, G. L., & Bullock, R. H. (Eds). (1987). Seeing for ourselves: Case-study research by teachers of writing. Portsmouth, NH: Heinemann.

Bogdan, R. C., & Biklen, S. K. (1992). Qualitative research for education: An introduction to theory and methods (2nd ed.). Boston: Allyn and Bacon.

Bogdan, R. C., & Biklen, S. K. (1998). Qualitative research for education: An introduction to theory and methods (3rd ed.). Boston: Allyn and Bacon.

Born, M. E. (1970). Psychological adaptation and development under acculturation stress: Toward a general model. Social Science and Medicine, 3, 529-547.

Bruck, M. (1978). The suitability of early French immersion programs for the language disabled child. The Canadian Modern Language Review, 34, 884-887.

Bruck, M. (1982). Language impaired children's performance in an additive education program. Applied Psycholinguistics, 3, 45-60.

Bruck, M. (1985a). Predictors of transfer out of early French immersion programs. Applied Psycholinguistics, 6, 39-61.

Bruck, M. (1985b). Consequences of transfer out of early French immersion programs. Applied Psycholinguistics, 6, 101-120.

Bullock, R. H. (1987). A quiet revolution: The power of teacher research. In G. L. Bissex & R. H. Bullock (Eds.), Seeing for ourselves: Case study research by teachers of writing (pp. 21-27). Portsmouth, NH: Heinemann.

Burke, C. (1980). The reading interview: 1977. In B. P. Farr & D. J. Stickler (Eds.). Reading comprehension: Resource Guide. Bloomington: School of Education, Indiana University.

Butler, K. G. (1997). Using dynamic approaches to support intervention: New approaches. In N. W. Nelson & B. B. Hoskins (Eds.), Strategies for supporting classroom success: Focus on communication- An audio workshop (pp. 1-8). San Diego, CA: Singular Publishing Company.

Campione, J., & Brown, A. (1985). Dynamic assessment: One approach and some initial data. (ERIC Document Reproduction Service No. ED 269 735).

Catts, H. W. (1997). The early identification of language-based reading disabilities. Language, Speech, and Hearing Services in Schools, 28(1), 86-89.

Catts, H. W. (1999, February). Language basis of reading difficulties: Implications for early identification and remediation. Springfield, MA.

Catts, H. W., & Kamhi, A. G. (1986). The linguistic basis of reading disorders: Implications for the speech-language pathologist. Language, Speech, and Hearing Services in Schools, 17, 329-341.

Catts, H. W., & Kamhi, A. G. (1999). Language and reading disabilities. Boston: Allyn and Bacon.

Chavez, D. V., Moran, V. R., Reid, S. L., & Lopez, M. (1997). Acculturation stress in children: A modification of the SAFE Scale. Hispanic Journal of Behavioral Sciences, 19(1), 34-44.

Clay, M. M. (1993). An observation survey of early literacy achievement. Portsmouth, NH: Heinemann.

Cloud, N. (1991). The acculturation of ethnolinguistic minorities. In A. Ambert (Ed.), Bilingual education and English as a second language: A research handbook, 1980-1990 (pp. 55-96). New York: Garland Press.

Cochran-Smith, M. (1991). Learning to teach against the grain. Harvard Educational Review, 61, 279-310.

Cochran-Smith, M., & Lytle, S. L. (1993). Inside/outside: Teacher research and knowledge. New York: Teachers College Press.

Cole, L. (1989, September). E pluribus pluribus: Multicultural imperative for the 1990s and beyond. ASHA, 65-70.

Cole, M, & Means, B. (1981). Comparative studies of how people think. Cambridge, MA: Harvard University Press.

Collier, C. (1988). Assessing minority students with behavioral problems. Linsdale, TX: Hamilton Publications.

Collier, V. (1987). Age and rate of acquisition of a second language for academic purposes, TESOL Quarterly, 21(4), 617-641.

Cuellar, I. Arnald, B., & Maldonado, R. (1995). Acculturation Rating Scale for Mexican Americans-II: A revision of the original ARSMA. Hispanic Journal of Behavioral Sciences, 17(3), 275-304.

Cuellar, I., Harris, L., & Jasso, R. (1980). An acculturation scale for Mexican American normal and clinical populations. Hispanic Journal of Behavioral Sciences, 2(3), 199-217.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. Review of Educational Research, 49, 222-251.

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In D. P. Dolson (Ed.), Schooling and language minority students: A theoretical framework (pp. 3-49). Los Angeles Evaluation, Dissemination, and Assessment Center. California State University.

Cummins, J. (1984). Bilingualism and special education: Issues in assessment and pedagogy. San Diego, CA: College-Hill Press.

Cummins, J. (1986). Empowering minority students: A framework for intervention. Harvard Educational Review, 56(1), 18-36.

Cummins, J. (1989). A theoretical framework for bilingual special education. Exceptional Children, 56(2), 111-119.

Damico, J. S. (1991). Descriptive assessment of communicative ability in limited English proficient children. In E. V. Hamayan & J. S. Damico (Eds.), Limiting bias in the assessment of bilingual students (pp. 157-219). Austin, TX: Pro-Ed.

Damico, J. S., & Hamayan, E. V. (1992). Multicultural language intervention: Addressing cultural and linguistic diversity. Chicago: Educom-Riverside.

DeFina, A. (1992). Portfolio assessment: Getting started. New York: Scholastic Professional Books.

Duran, R. H. (1989). Assessment and instruction of at-risk Hispanic students. Teaching Exceptional Children, 56(2), 145-152.

Edwards, S. A. (1997). A writing box for every child: Changing strategies for teaching writing in the first and second grade. Unpublished doctoral dissertation. University of Massachusetts, Amherst.

Escamilla, K., Andrade, K., Basurto, A. G., & Ruiz, O. A. (with Clay, M. M.). (1996). Instrumento de observacion de los logros de la lecto-escritura inicial. Portsmouth, NH: Heinemann.

Evans, J. E. (1996). Plotting the complexities of language sample analysis: Linear and dynamical models of assessment. In K. N. Cole, P. S. Dale, & D. J. Thal (Eds.). Assessment of communication and language (pp. 207-255). Baltimore: Paul H. Brooks Publishing Company.

Evans, J., & Craig, H. (1992). Language sampling collection and analysis: Interview compared to free play assessment contexts. Journal of Speech and Hearing Research, 35, 343-353.

Feuerstein, R. (1979). The dynamic assessment of retarded performers: The Learning Potential Assessment Device, theory, instruments, and techniques. Baltimore: University Park Press.

Feuerstein, R., Rand, Y., Haywood, H. C., Kyram, L., & Hoffman, M. B. (1995). Learning Propensity Assessment Device: New Experimental version. The International Center for the Enhancement of Learning Potential.

Figueroa, R. (1989). Psychological testing of linguistic minority students: Knowledge gaps and regulations. Exceptional Children, 56, 145-152.

Figueroa, R. A., & Garcia, E. (1994, Fall). Issues in testing students from culturally and linguistically diverse backgrounds. Multicultural Education, 10-19.

Flower, R. M. (1984). Delivery of speech-language pathology and audiology services. Baltimore: Williams and Wilkins.

Fountas, I. C., & Pinnell, G. S. (1996). Guided reading: Good first teaching for all children. Portsmouth, NH: Heinemann.

Fradd, S. H., & McGees, S. L. (with Wilen, D. K.). (1994). Instructional assessment: An integrative approach to evaluating student performance. Reading, MA: Addison-Wesley Publishing Company.

Gavillan-Torres, E. (1984). Issues of assessment of limited English proficient students and the truly disabled in the United Sates. In N. Miller (Ed.), Bilingualism and language disability: Assessment and remediation (pp. 131-153). San Diego, CA: College-Hill Press.

Gillam, R. B., & Carlile, R. M. (1997). Oral reading and story retelling of students with specific language impairments. Language, Speech, and Hearing Services in Schools, 28(1), 30-42.

Gillam, R.B., & McFadden, T. U. (1994). Redefining assessment as a holistic discovery process. Journal of Childhood Communication Disorders, 16(1), 36-40.

Gillam, R. B., Pena, E. D., & Miller, L. (1999). Dynamic assessment of narrative and expository discourse. Topics in Language Disorders, 20(1), 33-47.

Glazer, B. G., & Strauss, A. L. (1967). The discovery of grounded theory: Strategies for qualitative research. New York: Aldine.

Goddard School of Science and Technology Faculty. (1996). Student Writing Rubrics. Worcester, MA: Author.

Gonzales, M. D., & Kayser, H. (1997, November). Assessment of speech-language disorders in Hispanic individuals. Paper presented at the American Speech-Language-Hearing Association National Convention, Boston, MA.

Goodman, Y. M., & Burke, C. L. (1973). Theoretically based studies of patterns in oral reading performance. (Grant No. OEG-0-9-320375-4260). Washington, DC. U. S. Department of Health, Education, and Welfare.

Goodman, Y. M., & Goodman, K. S. (1998) To err is human: Learning about language. In C. Weaver (Ed.), Reconsidering a balanced approach to reading (pp. 101-123). Urbana, IL: National Council of Teachers of English.

Goodman, Y. M., Watson, D. J., & Burke, C. L. (1987). Reading miscue survey: Alternative assessment procedures. New York: Richard C. Owen Publishers.

Graves, T. (1967). Acculturation, access, and alcohol in a tri-ethnic community. American Anthropologist, 69, 306-321.

Grosjean, F. (1982). Life with two languages: An introduction to bilingualism. Cambridge, MA: Harvard University Press.

Gutierrez-Clellen, V. F. (1996). Language and diversity: Implications for assessment. In K. N. Cole, P. S. Dale, & Thal (Eds), Assessment of communication and language (pp. 29-56). Baltimore: Paul H. Brooks Publishing Company.

Gutierrez-Clellen, V. F. (1998). Syntactic skills of Spanish-speaking children with low school achievement. Language, Speech, and Hearing Services in Schools, 29(4), 207-215.

Gutierrez-Clellen, V. F. (1999). Language choice in intervention with bilingual children. American Journal of Speech-Language Pathology, 8(4), 291-302.

Gutierrez-Clellen, V. F., & Heinrichs-Ramos, L. (1993). Referential cohesion in the narratives of Spanish-speaking children: A developmental study. Journal of Speech and Hearing Research, 36, 559-567.

Gutierrez-Clellen, V. F., & Hofstetter, R. (1994). Syntactic complexity in Spanish narratives: A developmental study. Journal of Speech and Hearing Research, 37, 645-654.

Gutierrez-Clellen, V. F., Pena, E., & Quinn, R. (1995). Accommodating differences in narrative styles: A multicultural perspective. Topics in Language Disorders, 15(4), 54-67.

Gutierrez-Clellen, V. F., & Quinn, R. (1993). Assessing narratives of children from diverse cultural/linguistic groups. Language, Speech, and Hearing Services in Schools, 24, 2-7.

Gutierrez-Clellen, V. F., Restrepo, M. A., Bedore, L., Pena, E., & Anderson, R. (2000). Language sample analysis in Spanish-speaking children: Methodological considerations. Language, Speech, and Hearing Services in the Schools, 31(1), 88-98.

Hamayan, E. V., & Damico, J. S. (1991). Limiting bias in the assessment of bilingual students. Austin, TX: Pro-Ed.

Haywood, H. C., Brown, A. L., & Wingenfeld, S. (1990). Dynamic approaches to psychoeducational assessment. School Psychology Review, 19, 411-422.

Hiebert, E., Valencia, S., & Afflerbach, P. (1994). Definitions and perspectives. In S. Valencia, & P. Afflerbach. Authentic reading assessment: Practices and possibilities (pp. 6-23). Newark, DE: International Reading Association.

Hoover, J. J., & Collier, C. (1986). Classroom management through curricular adaptations: Educating minority students. Linsdale, TX: Hamilton Publishing Company.

Huang, R., Hopkins, J., & Nippold, M. A. (1997). Satisfaction with standardized language testing: A survey of speech-language pathologists. Language, Speech, and Hearing Services in Schools, 22, 236-241.

Hudelson-Lopez, S. (1977). Children's use of contextual clues in reading Spanish. Reading Teacher, 735-739.

Hunt, K. W. (1965). Grammatical structures written at three grade levels (Research Report No. 3). Champaign, IL: National Council of Teachers of English.

Jett-Simpson, M., & Leslie, L. (1997). Observing reading: The running record and comprehension. In L. Leshe & M. Jett-Simpson. Authentic literacy assessment: An ecological approach (pp. 96-115). New York: Longman.

Johnson, N. M. (1982). Assessment paradigm and atypical infants: An interventionist's perspective. In D. Bricker (Ed.), Intervention with at-risk and handicapped infants: From research to application (pp. 63-76). Baltimore: University Park Press.

Kamhi, A. G., & Catts, H. W. (1986). Toward an understanding of developmental language and reading disorders. Journal of Speech and Hearing Disorders, 51, 337-347.

Kayser, H. (1989). Speech and language assessment in Spanish-English speaking children. Language, Speech, and Hearing Services in Schools, 20, 226-244.

Kayser, H. (1995). Assessment of speech and language impairments in bilingual children. In H. Kayser (Ed.), Bilingual speech-language pathology: An Hispanic focus (pp. 185-206). San Diego, CA: Singular Publishing Company.

Kayser, H. (1998). Assessment and intervention resources for Hispanic children. San Diego, CA: Singular Publishing Company.

Kayser, H. G., Brice, A., Munoz, E., Moss, S. E., & Davis-McFarland, E. (1997, November). Research with culturally and linguistically diverse populations. A short course presented at the Annual American Speech-Language-Hearing Association National Convention, Boston, MA.

Kayser, H., & Restrepo, M. A. (1995). Language samples: Elicitation and analysis. In H. Kayser (Ed.) Bilingual speech-language pathology: An Hispanic focus (pp. 265-288). San Diego: Singular Publishing Group.

Keats, E. J. (1976). The snowy day. New York: Penguin Putnam Books for Young Readers.

Kiernan, B., & Swisher, L. (1990). The initial reading of novel English words: Two single subject experiments with minority-language children. Journal of Speech and Hearing Research, 33, 707-716.

Knapp, M. (1979). Ethnographic contributions to evaluation research: The experimental schools program evaluation and some alternatives. In T. D. Cook & C. S. Reichardt (Eds.), Qualitative and quantitative methods in evaluation research (pp. 118-139). London: Sage Publications.

Kratcoski, A. M. (1998). Guidelines in using portfolios in assessment and evaluation. Language, Speech, and Hearing Services in Schools, 29(1), 3-10.

Langdon, H. W. (1989). Language disorder or difference? Assessing the language skills of Hispanic students. Exceptional Children, 56, 160-167.

Langdon, H. W. (1992). Speech and language assessment of LEP/bilingual Hispanic children. In H. W. Langdon (with L. R. L. Cheng) (Eds.), Hispanic children and adults with communication disorders: Assessment and intervention (pp. 201-271). Gaithersburg, MD; Aspen.

Langdon, H. W., & Saenz, T. I. (1996). Speech-language assessment of bilingual/bicultural students. In H. W. Langdon & T. I. Saenz (Eds.), Language assessment and intervention with bilingual students: A guide for speech-language-hearing professionals (pp. 3-30). Oceanside, CA: Academic Communication Associates.

Larson, V., & McKinley, N. (1995). Language disorders in older students. Eau Claire, WI: Thinking Publications.

Launer, P. (1998, June). Get real! Authentic assessment and intervention. Paper presented at the Fifteenth Annual Language and Learning Disabilities Institute, Emerson College, Boston, MA.

Leshe, L., & Jett-Simpson, M. (1997). Authentic literacy assessment: An ecological approach. New York: Longman.

Leslie, L., & Dussault, N. F. (1997). Issues of reliability and validity. In L. Leshe & M. Jett-Simpson (Eds.), Authentic assessment: An ecological approach (pp. 232-243). New York: Longman.

Lidz, C. S. (Ed.). (1987). Dynamic assessment: An interactional approach to evaluating learning potential. New York: Guilford Press.

Lidz, C. S. (1991). Practitioners guide to dynamic assessment. New York: Guilford Press.

Lidz, C. S., & Pena, E. D. (1996). Dynamic assessment: The model, its relevance as a nonbiased approach, and its application to Latino American preschool children. Language, Speech, and Hearing Services in Schools, 27(4), 367-372.

Lidz, C. S., & Thomas, C. (1987). The preschool learning assessment device: Extension of a static approach. In C. S. Lidz (Ed.), Dynamic assessment: An interactional approach to evaluating learning potential (pp. 288-326). New York: Guilford Press.

Linares, N., & Sanders, L. (1977). Evaluation of syntax in three-year old Spanish-speaking Puerto Rican children. Journal of Speech and Hearing Research, 20, 350-357.

Linder, T. W. (1993). Transdisciplinary play-based assessment: A functional approach to working with young children (rev. ed.). Baltimore: Paul H. Brooks Publishing Company.

Lopez, E. C., Lamar, D., & Scully-Demartini, D. (1997). The cognitive assessment of limited English-proficient children: Current approaches and partial recommendations. Cultural Diversity and Mental Health, 3(2), 117-130.

Maldonado, J. A. (1994). Bilingual special education: Specific learning disabilities in language and reading. The Journal of Educational Issues of Language Minority Students, 14, 127-148.

Marin, B. V., Otero-Sabogal, & Perez-Stable, E. J. (1987). Development of a Short Acculturation Scale for Hispanics. Hispanic Journal of Behavioral Sciences, 9(2), 183-205.

Martinez, R., Norman, R. D., & Delaney, H. (1984). A Children's Hispanic Background Scale. Hispanic Journal of Behavioral Sciences, 6(2), 103-112.

Massachusetts Department of Education. (2000). Report of 1999 Massachusetts and local school districts' MCAS results by race/ethnicity. [on-line], Available: http://www.doe.mass.edu/mcas.

Massachusetts English Language Assessment-Oral (1994). Arlington, VA: Evaluation Assistance Center East, George Washington University.

Mattes, L. J., & Omark, D. R. (1984). Speech and language assessment for the bilingual handicapped. San Diego, CA: College-Hill Press.

Maxwell, M. M. (1990). The authenticity of ethnographic research. Journal of Childhood Communication Disorders, 13(1), 1-12.

Mayer, M. (1975). One Frog Too Many. New York: Dial Press.

Mayer, M. (1995). Just Grandma and Me. Cambridge, MA: Broderbund Software.

McCauley, R., & Swisher, L. (1984). Use and misuse of norm-referenced tests in clinical assessment: A hypothetical case. Journal of Speech and Hearing Disorders, 49, 338-348.

McLaughlin, B., Blanchard, A. G., & Osanai, Y. (1995, Summer). Assessment in bilingual preschool children. National Clearinghouse for Bilingual Program Information Series, Number 22.

Mehan, H., Hertwreck, H., & Meihls, J. L. (1986). Handicapping the handicapped. Stanford, CA: Stanford University Press.

Merino, B. J. (1992). Acquisition of syntactic and phonological features in Spanish. In H. W. Langdon & L. L. Cheng (Eds.). Hispanic children and adults with communication disorders: Assessment and intervention. Gaithersburg, MD: Aspen.

Merrell, A. W., & Plante, E. (1997). Norm-referenced test interpretation in the diagnostic process. Language, Speech, and Hearing Services in Schools, 28(1), 50-58.

Mertens, D. M. (1998). Research methods in education and psychology: Integrating diversity with qualitative and quantitative approaches. Thousand Oaks, CA: Sage Publications.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement: Third edition (pp. 13-104). New York: American Council of Education, MacMillan Publishing Company's National Council on Measurement in Education.

Miriam, S. B. (1998). Qualitative research and case study applications in education. San Francisco, CA: Jossey-Boss Publishers.

Mitchell, R. (1992). Testing for learning. New York: The Free Press.

Montgomery, G. T. (1992). Comfort and acculturation status among students from South Texas. Hispanic Journal of Behavioral Sciences, 14(2), 201-223.

Morrow, L. M. (1988). Retelling stories as a diagnostic tool. In S. M. Glazer, L. W. Searfoss, & L. M. Gentile (Eds.). Re-examining reading diagnosis: New trends and procedures (pp. 128-149). Newark, DE: International Reading Association.

National Clearinghouse for Bilingual Education. (1997, October). High-stakes assessment: A research agenda for English-language learners. [on-line], Available: http://www.ncbe.gwu.edu

Navarrete, C., & Gustke, C. (1996). A guide to performance assessment for linguistically diverse students. National Clearinghouse for Bilingual Education. [on-line] Available : http:www.ncbe.edu/miscpubs/eacwest/perform.htl

Navarrete, C., Wilde, J., Nelson, C., Martinez, R., & Hargett, G. (1990, Summer). Informal assessment in educational evaluation: Implications for bilingual education programs. National Clearinghouse for Bilingual Education Program Information Guide Series, Number 3.

Nelson, N. W. (1989). Curriculum-based language assessment. Language, Speech, and Hearing Services in Schools, 20(2), 170-184.

Nelson, N. W. (1992). Targets of curriculum-based language assessment. Best Practices in School Speech-Language Pathology, 2, 73-86.

Nelson, N. W. (1993). Childhood language disorders in context: Infancy through adolescence. New York: Merrill/MacMillian.

Nelson, N. W. (1994). Curriculum-based language assessment across the grades. In G. P. Wallach & K. G. Butler (Eds.), Language learning disabilities in school-age children and adolescents: Some principles and applications (pp. 104-131). New York: MacMillian College Publishing Company.

Nelson, N. W. (1998). Childhood language disorders in context: Infancy through adolescence (2nd ed.). Boston: Allyn and Bacon.

Nicolosi, L., Harryman, E., & Kresheck, J. (1978). Terminology of communication disorders: Speech, language, hearing. Baltimore: Williams and Wilkins Company.

Notari-Syverson, A., & Losardo, A. (1996). Assessing children's language in meaningful contexts. In K. N. Cole, P. S. Dale, & D. J. Hall (Eds.), Assessment of communication and language (pp. 257-279). Baltimore: Paul H. Brooks Publishing Company.

Ochoa, S. H., Galarza, A., & Gonzalez, D. (1996). An investigation of school psychologists' assessment practices of language proficiency with bilingual and limited English-proficient students. Diagnostique, 21(4), 17-36.

Office for Civil Rights, U. S. Department of Education. (1993). The Fall 1990 Elementary and Secondary School Civil Rights Survey. Washington, DC: Author.

Olmedo, E. (1981). Testing linguistic minorities. American Psychologist, 36(10), 1078-1085.

Olmedo, E. L. (1980). Quantitative models of acculturation: An overview. In A. M. Padilla (Ed.), Acculturation: Theory, model, and some new findings (pp. 27-46). Boulder, CO: Westview Press.

Olmedo, E. L., Martinez, J. L., & Martinez, S. R. (1978). A measure of acculturation for Chicano adolescents. Psychological Reports, 42, 159-170.

O'Malley, J. M. (1996). Authentic assessment for English language learners: Practical approaches for teachers. Reading, MA: Addison Wesley Publishing Company.

Ortiz, A. A. (1984). Choosing the language of instruction for exceptional bilingual children. Teaching Exceptional Children, 16, 208-212.

Ortiz, A. A., & Garcia, S. B. (1990). Using language assessment data for language and instructional planning for bilingual exceptional children. In A. L. Carrasquillo & R. E. Baecher (Eds.), Teaching the bilingual special education student (pp. 25-47). Norwood, NJ: Ablex Publishing Company.

Ortiz, A. A., & Polyzoi, E. (1988). Language assessment of Hispanic learning disabled and speech and language handicapped students: Research in progress. In A. A. Ortiz & B. A. Ramirez (Eds.), Schools and the culturally diverse exceptional student: Promising practices and future directions. Reston, VA: The Council for Exceptional Children.

Padilla, A. M. (1980). The role of cultural awareness and ethnic loyalty in acculturation. In A. M. Padilla (Ed.), Acculturation: Theory, models, and some new findings (pp. 47-84). Boulder, CO: Westview Press.

Payan, R. (1984). Language assessment for bilingual exceptional children. In L. M. Baca & H. T. Cervantes (Eds.), The bilingual special education interface (pp. 125-137). St Louis: Times Mirror/Mosby.

Pena, E. D. (1993). Dynamic assessment: A nonbiased approach for assessing the language of young children. Unpublished doctoral dissertation, Temple University.

Pena, E. D. (1996). Dynamic assessment: The model and its language applications. In K. N. Cole, P. S. Dale, & D. J. Thal (Eds.), Assessment of communication and language (pp. 281-307). Baltimore: Paul H. Brooks Publishing Company.

Pena, E. D., Quinn, R., & Iglesias, A. (1992). The application of dynamic methods to language assessment: A nonbiased procedure. Journal of Special Education, 26(3), 269-280.

Perrozi, J. (1985). A pilot study of language facilitation for bilingual, language-handicapped children: Theoretical and clinical implications. Journal of Speech and Hearing Disorders, 50, 403-406.

Perrozi, J., & Chavez-Sanchez, M. L. (1992). The effects of instruction in L1 on receptive acquisition of L2 for bilingual children with language delays. Language, Speech, and Hearing Services in Schools, 23(4), 348-352.

Pierce, L. V., & O'Malley, J. M. (1992, Spring). Performance and portfolio assessments for language minority students. National Clearinghouse for Bilingual Education Program Information Guide Series, Number 9.

Pinnell, G. S., Pikluski, J. J., Wixson, K. K., Campbell, J. R., Gough, P. B., & Beatty, A. S. (1995). Listening to children read aloud: Data from NAEP's Integrated Reading Performance Record (RPR) at Grade 4. Report No. 23-FR-04, prepared by the Educational Testing Service. Washington, DC: Office of Educational Research and Improvement, U. S. Department of Education.

Plante, E., & Vance, R. (1994). Selection of preschool language tests: A data-based approach. Language, Speech, and Hearing Services in Schools, 25, 15-24.

Porch, B. (1967). The Porch Index of Communicative Ability. Palo Alto, CA: Consulting Psychologists Press.

Prelock, P. A. (1997). Language-based curriculum analysis: A collaborative assessment and intervention process. Journal of Children's Communication Development, 19(1), 35-42.

Quinn, R., Goldstein, B., & Pena, E. D. (1996). Cultural linguistic variation in the United States and its implications for assessment and intervention in speech-language pathology: An introduction. Language, Speech, and Hearing Services in Schools, 27(4), 345-346.

Restrepo, M. A. (1998). Identifiers of predominantly Spanish-speaking children with language impairments. Journal of Speech, Language, and Hearing Research, 41, 1398-1411.

Restrepo, M. A., & Gutierrez-Clellen, V. F. (1997, November). Data-based approach to the language assessment of Spanish-speaking children. Paper presented at the American Speech-Language-Hearing Association. National Conference, Boston, MA.

Rivera-Viera, D. T. (1978). A descriptive study of reading miscues of Spanish-speaking elementary school children. Unpublished doctoral dissertation, University of Massachusetts, Amherst.

Roseberry-McKibbin, C. (1995). Multicultural students with special needs. Oceanside, CA: Academic Communication Associates.

Roseberry-McKibbin, C. A., & Eicholtz, G. E. (1994). Serving children with limited English proficiency in the schools: A national survey. Language, Speech, and Hearing Services in Schools, 25(3), 156-164.

Routman, R. (1994). Invitations: Changing as teachers and learners K-12. Portsmouth, NH: Honeymoon.

Schiff-Myers, N. (1992). Considering arrested language development and language loss in the assessment of second language learners. Language, Speech, and Hearing Services in Schools, 23, 28-33.

Secord, W. A., Wiig, E. H., Damico, J. S., & Goodin, G. L. (1994). Classroom communication assessment: Evaluating performance in context. Chicago: The Riverside Publishing Company.

Schraeder, T., Quinn, M., Stockman, I. J., & Miller, J. (1999). Authentic assessment as an approach to preschool language screening. American Journal of Speech-Language Pathology, 8(3), 195-200.

Seidman, I. (1998). Interviewing as qualitative research: A guide for researchers in education and social science (2nd ed.). New York: Teachers College Press.

Semel, E. M., Wiig, E. H., & Secord, W. A. (1989). CELF-R: Screening. San Antonio, TX: The Psychological Corporation.

Sewell, T. E. (1987). Dynamic assessment as a nondiscriminatory procedure. In C. S. Lidz (Ed.), Dynamic assessment: An interactional approach to evaluating learning potential (pp. 426-443). New York: Guilford Press.

Shutnabb-Kangas, T. (1984). Bilingualism or not: The education of minorities. Clevedon, England: Multilingual Matters Ltd.

Shutnabb-Kangas, T., & Toukomaa, P. (1976). Teaching migrant children's mother tongue and learning the language of the host country in the context of the socio-cultural situation of the migrant family. Helsinki: The Finnish National Commission for UNESCO.

Silliman, E. R., & Wilkinson, L. C. (1991). Communicating for learning: Classroom observation and collaboration. Gaithersburg, MD: Aspen.

Singer, B. D., & Bashir, A. S. (1998, September). Written language intervention for students with language-learning disabilities: Enhancing executive functions, self-regulation, and language. Springfield, MA.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). Preventing reading difficulties in young children. Washington, DC: National Academy Press.

Snow, C. E., Scarborough, H. S., & Burns, M. S. (1999). What speech-language pathologists need to know about early reading. Topics in Language Disorders, 20(1), 48-58.

Spradley, J. P. (1979). The ethnographic interview. New York: Holt, Rinehart, and Winston.

Sticht, T. G., Beck, L., Hauke, R., Kleiman, G., & James, J. (1974). Auding and reading comprehension: A developmental model. Alexander, VA: Human Resources Research Organization.

Sticht, T. G., & James, J. (1984). Listening and reading. In P. D. Pearson (Ed.). Handbook of reading research (pp. 292-317). New York: Longman.

Stockman, I. (1996). The promises and pitfalls of language sampling analysis as an assessment tool for linguistic minority children, Language, Speech, and Hearing Services in Schools, 27(4), 355-366.

Success For All Foundation (1998). Lee conmigo: Participants Training Book. Baltimore: Author.

Sumner, H. (1991). Whole language assessment and evaluation: A special education perspective. In W. Harp (Ed.). Assessment and evaluation in whole language programs (pp. 147-168). Norwood, MA: Christopher Gordon Publishers.

Synder, S., & Downey, D. M. (1991). The language reading relationship in normal and learning disabled children. Journal of Speech and Hearing Research, 34, 129-140.

Taylor, O. (1986). Issues, historical perspectives, and conceptual framework. In O. Taylor (Ed.). Treatment of communication in culturally and linguistically diverse populations (pp. 3-19). San Diego, CA: College-Hill Press.

Taylor, O. (1993, March). Cultural diversity in a changing America: New directions in communication disorders. Paper presented at the Fourteenth Annual Communication Disorders Conference, Boston, MA.

Taylor, O., & Payne, K. (1983). Culturally valid testing: A proactive approach. Topics in Language Disorders, 3, 8-20.

Thordardottir, E. T., Weismer, S. E., & Smith, M. E. (1997). Vocabulary learning in bilingual and monolingual intervention. Child Language Teaching and Therapy, 13(3), 251-227.

Tzuriel, D., & Klein, P. (1987). Assessing the young child: Children's analogical thinking and modifiability. In C. S. Lidz (Ed.), Dynamic assessment: An interactional approach to evaluating learning potential (pp. 268-287). New York: Guilford Press.

Valencia, S., & Afflerbach, P. (1994). Authentic reading assessment: Practices and possibilities. Newark, DE: International Reading Association.

Valle-Arroyo, F. (1996). Dual route models in Spanish: Developmental and neuropsychological data. In M. Carreiras, J. E. Garcia-Albea, & N. Sebastian-Galles (Eds.). Language processing in Spanish (pp. 89-118). Mahwah, NJ: Lawrence Erlbaum Associates.

van Kleek, A. (1990). Emergent literacy: Learning about print before learning how to read. Topics in Language Disorders, 10(2), 25-45.

van Kleek, A. (1998). Preliteracy domains and stages: Laying the foundations for beginning reading. Journal of Children's Communication Development, 20(1), 33-51.

Vaughn-Cooke, F. B. (1983). Improving language assessment in minority children, ASHA, 25(9), 29-34.

Wallach, G. P., & Butler, K. G. (1994). Language learning disabilities in school-age children and adolescents: Some principles and applications. New York: MacMillian College Publishing Company.

Weaver, C. (1980). The other side of reading. Carthage, IL: Good Apple.

Weaver, C. (1994). Reading process and practice: From sociopsycholinguistics to whole language (2nd ed.). Portsmouth, NH: Heinemann.

Weiner, P., & Hoock, W. (1973). The standardization of tests: Criterion and criticisms. Journal of Speech and Hearing Research, 16, 616-626.

Weismantel, M. J., & Fradd, S. H. (1989). Understanding the need for change. In S. H. Fradd & M. J. Weismantel (Eds.), Meeting the needs of culturally and linguistically diverse students: A handbook for educators (pp. 1-13). Austin, TX: Pro-Ed.

Wells, G. (1986). The meaning makers: Children learning language and using language to learn. Portsmouth, NH: Heinemann.

Westby, C. E. (1980). Language abilities through play. Language, Speech, and Hearing Services in Schools, 21, 154-168.

Westby, C. E. (1990). Ethnographic interviewing: Asking the right questions to the right people in the right ways. Journal of Childhood Communication Disorders, 13(1), 101-112.

Westby, C. E. (1992). Narrative analysis. Best Practices in Speech-Language Pathology, 2, 53-64

Westby, C. E. (1994). The effects of culture on genre, structure, and style on oral and written texts. In G. P. Wallach & K. G. Butler (Eds.), Language learning disabilities in school-age children and adolescents: Some principles and applications (pp. 180-218). New York: MacMillian College Publishing Company.

Westby, C. E. (1996, October). Multicultural issues in speech-language pathology: Communication in a culturally diverse world. Boston, MA.

Westby, C. E., Van Dongen, P., & Maggart, Z. (1989). Assessing narrative competence. Seminars in Speech and Language, 10(1), 63-71.

Westman, J. C. (1990). Handbook of learning disabilities: A multisystematic approach. Boston: Allyn and Bacon.

Whitmire, K. (2000). Action: School services. Language, Speech, and Hearing Services in Schools, 31(2), 194-197.

Wiener, R. B., & Cohen, J. H. (1997). Literacy portfolios: Using assessment to guide instruction. Upper Saddle River, NJ: Merrill Publishing Company.

Wiggins, G. (1990, December). The case for authentic assessment. ERIC Clearinghouse on Tests and Measurements. Washington, DC: American Institute for Research.

Wiig, E. H., & Secord, W. A. (1994). Classroom and Communication Checklist. In W. A. Secord, E. H. Wiig, S. H. Coolahan, & D. H. Pallante. Team-based problem solving (pp. 37-39). Chicago, IL: Riverside Publishing Company.

Wiig, E. H., & Secord, W. A. (1991). Measurement and assessment: Making sense of test results. Chicago, IL: The Riverside Publishing Company.

Wiig, E. H., & Semel, E. (1984). Language assessment and intervention for the learning disabled (2nd ed.). Columbus, OH: Charles E. Merrill Publishing Company.

Williams, C. I., & Berry, J. W. (1994). Primary prevention of acculturation stress among refugees. American Psychologist, 46, 632-641.

Wolf, D., Bixby, J., Glenn, J., III, & Gardner, H. (1991). To use their minds well: Investigating new forms of student assessment. Review of Research in Education, 17, 31-74.

Worcester Public Schools (1995, March). Worcester Public Schools PreK-12 English Language Arts Curriculum. Worcester, MA: Author.

Worcester Public Schools Office of Staff and Program Development. (1998, Fall). Your window to opportunity: Staff Development Offerings-Supporting Programs. [Brochure}. Worcester, MA: Author.

Worcester Public Schools (1999). Curriculum-based reading and language arts inventory. Worcester, MA: Author.

Worcester Public Schools. (n.d.). Language Preference and Home Language Survey Form. Worcester, MA: Author.

Worthen, B. (1993, February). Critical issues that will determine the future of alternative assessment. Phi Delta Kappa, 444-454.

Zehler, A. M., Hopstock, P. J., Fleischman, H. J., & Greniuk, C. (1994). An examination of assessment of limited English proficient students. Task Order D070 Report, March 28, 1994, Arlington, VA: Special Issues Analysis Task Force Report.